Master's Thesis in Physics

Search for right-handed neutrinos with 79.8 fb^{-1} of data collected at $\sqrt{s} = 13$ TeV with the ATLAS detector

20th November, 2018

Author:
Helle Gormsen

Supervisor:
Assoc. Prof. Stefania Xella

University of Copenhagen
Niels Bohr Institute
High energy physics group

Abstract

A search for a heavy neutral lepton (HNL) of Majorana nature is performed using data collected by the ATLAS detector at $\sqrt{s} = 13$ TeV which corresponds to an integrated luminosity of 79.8 fb^{-1}. The properties of the HNL are described in the νMSM which presents a possible extension of the Standard Model. The search is conducted in the prompt $W^{\pm} \rightarrow \mu^{\pm}\mu^{\pm}e^{\mp}\nu_e$ channel for which the mixing angle $V_{\mu N}$ is probed, where $V_{\mu N}$ is the matrix element describing the mixing between the Standard Model neutrino of flavor μ and the heavy neutral lepton N. In order to accurately estimate the background processes which contain fake or non-prompt leptons, the data-driven technique called the fake factor method is applied. The estimated background is compared to data in control regions which proves to be consistent. The sensitivity of the search is estimated for five mass points of the heavy neutral lepton, $M_N = 5$, 10, 20, 30 and 50 GeV. For the largest mass points, the expected sensitivity of the search is well below the current state-of-the-art limits set by the CMS and DELPHI experiments.

Acknowledgements

I would first like to thank my thesis advisor Stefania Xella for her great guidance. She has consistently steered me in the right direction for my research and writing and provided me valuable feedback. I would also like to thank Hass Abouzeid and Fabian Thiele who have helped me with many technical problems and given me advice for how the analysis can be performed and improved. Finally, I am incredibly grateful to my parents and Kieran Dawson for their support and continuous encouragement throughout the process. This accomplishment would not have been possible without you. Thank you.

Contents

1 Introduction

The Standard Model (SM) [1] of particle physics has proved successful in describing many phenomena, however several observations within particle physics, astrophysics and cosmology remain unexplained. Three of the confirmed beyond the SM (BSM) mysteries are: the nature of dark matter; the masses of the neutrinos and their small size; and the baryon asymmetry of the universe. Many theories have been suggested with possible extensions of the SM among which is the Neutrino Minimal Standard Model (νMSM) [2–4]. By adding three heavy right-handed Majorana neutrinos to the SM, heavy neutral leptons (HNL), the model can explain all the confirmed BSM phenomena without introducing new energy scales. In the νMSM, the very small neutrino masses are described via the see-saw mechanism, a dark-matter candidate is provided by the lightest of the three HNLs, and the baryon asymmetry in the universe is explained given the two additional HNLs are mass-degenerate. The requirements imposed by the experimental observations are used as bounds for the model parameters and thereby defines where the HNLs can be found[5].

The right-handed neutrinos are produced through mixing with the SM neutrinos for which the coupling is governed by the mixing angle V_{lN}, where l is the flavor of the SM neutrino and N is the right-handed neutrino. This is a rare process so it is crucial to both have a lot of data and an efficient signal selection. Previously, limits on the νMSM parameters have been set by e^+e^- collider experiments as the Large Electron-Positron Collider (LEP) at CERN and fixed-target experiments [6]. Due to the increasing luminosity at the LHC and its high intensity collisions, it provides great conditions to further probe the parameter space of the νMSM. Depending on the mass of the HNLs and their coupling to the SM neutrinos, they can travel some distance in the detector before decaying or have a prompt decay after production.

In this thesis, a search for HNLs is performed using the ATLAS detector at the CERN LHC [7]. The data is collected in 2015-2017 at a centre-of-mass energy of $\sqrt{s}$ and corresponds to an integrated luminosity of 79.8 fb^{-1}. The HNL mass is restricted to be below the mass of the W-boson so the very high rate of W-boson production at the LHC is exploited. The search is conducted in the prompt $W^\pm \to \mu^\pm\mu^\pm e^\mp\nu_e$ channel for which the mixing angle $V_{\mu N}$ is probed. This channel exploits the Majorana property of the right-handed neutrinos which makes it possible to conduct the trilepton search in a final state with no opposite-sign same-flavor lepton pairs.

Due to the rare topology of the considered HNL final state, very few background processes mimic the signal. This makes the ability to accurately estimate background processes involving fake or non-prompt leptons crucial for the sensitivity of the search. A data-driven analysis is used to determine the fake background in the analysis using the so-called fake factor method. The stability of the data-driven model is checked in regions which are similar but independent to the signal region and subsequently applied in the signal region where the expected discovery significance for the signal samples is determined.

The thesis is organized as follows. In Sec. 2, the SM is briefly presented and the properties of the νMSM are examined. The parameter restrictions which are relevant for the scope of the analysis are presented. Section 3 focuses on the signal channel, the related cross section and the implementation of the right-handed neutrino in the SM process. The considered HNL masses and their respective decay length and mixing angle with the SM neutrino are listed. In Sec. 4, the LHC and the ATLAS detector are introduced. The applied data and MC samples are presented in Sec. 5. Section 6 presents the basic concepts for defining the analysis procedure. The selection requirements, which are applied in the search for prompt HNL signal events, are defined in Sec. 7. The optimal requirements for the signal region are chosen by maximizing the expected discovery significance. Section 8 is devoted to the background estimation method. The data-driven model is developed and control regions are defined which are used to check the stability of the model. Section 9 focuses on the systematic uncertainties related to the applied fake factor method. Finally, the background estimation model is used in the signal region in Sec. 10 and the expected sensitivity to the mixing angle is presented.

2 Theory

The existence of right-handed neutrinos is predicted by the νMSM which introduces three heavy right-handed Majorana neutrinos. In order for the theory to explain observed BSM phenomena, the right-handed neutrinos must have a mass below the electroweak scale [4] which means that they can produced at the LHC at some rate. The search for sterile neutrinos in data collected by experiments at the LHC is thus strongly motivated by theorists.

In this section, a brief introduction to the theory behind the νMSM is presented. First, I am giving an overview of the SM. Thereafter, the νMSM is examined focusing on the neutrino mass term and the see-saw mechanism. Finally, the restrictions on the free parameters of the νMSM, which are set by observations of BSM phenomena, are studied.

2.1 The Standard Model

The SM is an incredibly successful theory of the fundamental constituents of the universe: the elementary particles and their interactions [1]. It is formulated as a renormalizable Quantum Field Theory (QFT) which is based on the $SU(3) \times SU(2) \times U(1)$ gauge group. In the SM, twelve fundamental particles called fermions are responsible for the matter content of the universe. The fermions interact through four fundamental forces: the electromagnetic (EM) force described by Quantum Electrodynamics (QED); the strong force described by Quantum Chromodynamics (QCD); the weak force described by the Electroweak theory; and gravity which is responsible for large scale dynamics. The three forces which are relevant for particle physics (strong, EM, and weak) are all described by a QFT and correspond to the exchange of a spin-1 gauge boson which is the force-carrying particle. The final element of the SM is the Higgs boson which was discovered by the ATLAS and CMS experiments at the LHC in 2012 [8] [9]. Unlike the other particles, the Higgs boson is a spin-0 scalar particle. In the SM, it is responsible for the mechanism through which all the other particles acquire mass.

The fermions in the SM are divided into two categories: leptons and quarks. In Tab. 1, properties of the twelve fermions are listed. Each fermion has an anti-particle with an equal but opposite charge and they are all spin-1/2 particles. In the table, it is also listed which forces affect each fermion. The strong force interacts only with quarks which have colour quantum numbers, the EM force interacts with all charged particles, and the weak force interacts with all fermions. When the SM was formulated in the 1960's, neutrinos were thought to be massless and lepton numbers conserved. In the SM, they are the only particles described by just left-handed Weyl spinors. The mass terms for quarks and charged leptons are generated by their Yukawa coupling to the Higgs field for which both a left-handed and right-handed fermion field is necessary. Through observations of neutrino oscillations [10] it was found that neutrinos have a small but non-zero charge. The complications of the observed neutrino masses and further BSM phenomena will be discussed in the next section.

In Tab. 2, the four fundamental forces and their properties are listed. Gravity is included for consistency even though gravity can not be included in a renormalizable QFT yet [12]. At small scales, the strong force is dominating for quarks. Due to the strong force, quarks or gluons are never observed as free particles but confined to bound states called hadrons. The EM interactions are mediated by the exchange of virtual photons, and the weak charged-current and neutral-current interactions by the W and Z-bosons respectively.

2.2 The Neutrino Minimal Standard Model

With the observation of neutrino masses and lepton flavour violation in neutrino experiments [10], it is clear that the SM is not a complete theory. Also the existence of dark matter and dark energy and the observed baryon asymmetry in the universe are observed phenomena within particle physics, astrophysics and cosmology which the SM is not able to explain [4]. By extending the SM

Generation	Particle	Charge	Mass/MeV	Interactions
		Leptons		
First	Electron: e^-	-1	0.5109	EM, Weak
	Neutrino: ν_e	0	$< 2 \cdot 10^{-6}$	Weak
Second	Muon: μ^-	-1	105.66	EM, Weak
	Neutrino: ν_μ	0	$< 2 \cdot 10^{-6}$	Weak
Third	Tau: τ^-	-1	1776.86	EM, Weak
	Neutrino: ν_τ	0	$< 2 \cdot 10^{-6}$	Weak
		Quarks		
First	Down: d	-1/3	4.7	Strong, EM, Weak
	Up: u	+2/3	2.2	Strong, EM, Weak
Second	Strange: d	-1/3	96	Strong, EM, Weak
	Charm: c	+2/3	$1.28 \cdot 10^3$	Strong, EM, Weak
Third	Bottom: b	-1/3	$4.18 \cdot 10^3$	Strong, EM, Weak
	Top: t	+2/3	$173.1 \cdot 10^3$	Strong, EM, Weak

Table 1: The twelve fermions in the SM which are the known elementary particles [1]. All fermions have spin 1/2. The possible interactions of each particle are governed by the forces experienced by the particle. The listed values are taken from the PDG [11].

Force	Strength	Boson	Spin	Mass/GeV
Strong	1	Gluon: g	1	0
Electromagnetism	10^{-3}	Photon: γ	1	0
Weak	10^{-8}	W-boson: $W^\pm$	1	80.39
		Z-boson: Z	1	91.19
Gravity	10^{-37}	Graviton?: G	2	0

Table 2: The four known forces of nature: strong, electromagnetic, weak, and gravity. The role of the force carrier for gravity, popularly referred to as gravitons, is not well known. Relative strengths of the forces are listed which are approximate values for two fundamental particles which are separated 1 fm=10^{-15} m [1]. The listed masses are taken from the PDG [11].

with three gauge-singlet neutrinos with masses smaller than the electroweak scale, it is possible to achieve a theory which can theoretically predict the observed BSM phenomena [13]. These neutrinos are referred to as being *singlet* or *sterile* since they have zero strong, EM and weak charges and thus do not interact with the gauge bosons listed in Tab. 2. In order to theoretically describe the measurements of the neutrino oscillations, it is not sufficient to just generate a mass term for neutrinos; the model also has to account for the extremely small size of the neutrino masses. Therefore, it is necessary for the introduced right-handed neutrinos to have Majorana properties and thereby allow for a see-saw mechanism to explain the size of the neutrino mass.

The considered theory is called the νMSM as mentioned above. It is a strongly motivated and popular theory due to its clear applicability to describe neutrino flavour oscillations, the simplicity of the SM extension, and its ability to explain the observed BSM phenomena without introducing new energy scales [4]. As mentioned above, the active neutrinos in the SM are described by left-handed Weyl spinors which do not have an associated mass term. They are referred to as *active* since they interact through the weak force in contrast to the right-handed neutrinos which are sterile. When extending the SM with right-handed neutrino fields, mass terms are generated. The foundation of the νMSM is studied in the next section where focus will be on the different fermion fields which are solutions to the Dirac equation and the way mass terms are generated.

2.2.1 Dirac, Weyl and Majorana fermions

In order to describe the effect of extending the SM with right-handed neutrinos, some of the basic properties of the SM are first presented. The SM is a relativistic formulation of quantum mechanics for which the Dirac equation is used to describe the fundamental fermions [12]. The Dirac equation is defined as:

$$(i\gamma^\mu \partial_\mu - m)\Psi = 0, \tag{2.1}$$

for which Einstein notation is used to indicate a sum over the four indices of μ. The solution to the Dirac equation can be expressed as both a Dirac, Weyl and Majorana field. A Dirac field is the unconstrained solution to the Dirac equation which is represented by a four-component spinor whereas both Weyl and Majorana fermions are simpler solutions. The Weyl field is massless and the Majorana field does not change under charge conjugation so Majorana particles are their own anti-particles. In the SM, the massive leptons and all quarks are represented by Dirac fields whereas the massless neutrinos are represented by Weyl fields. No Majorana fields are a part of the SM since it would break the gauge invariance of the SM. In the following, I will refer to a Dirac field using Ψ, a Weyl field using ψ, and a Majorana field using χ. The dynamics of the fermion fields are governed by the Lagrangian density which for Dirac fields is, [12]

$$\mathcal{L}_D = i\bar{\Psi}\gamma^\mu \partial_\mu \Psi - m\bar{\Psi}\Psi, \tag{2.2}$$

where the first term and second term represent the kinetic and potential energy density of the field respectively. The field $\bar{\Psi}$ refers to the Dirac adjoint of the field defined as: $\bar{\Psi} = \Psi^\dagger \gamma^0$. The variable m is the mass of the massive fermion field which is a function of the field's Yukawa coupling to the Higgs field and the vacuum expectation value of the Higgs field [12]. For Weyl and Majorana fields, the Lagrangian can be simplified. First I will focus on the properties of the Weyl field. Weyl fermions are massless particles described by a two-component spinor. The Dirac field can be expressed in terms of two Weyl fiels:

$$\Psi = P_L\Psi + P_R\Psi = \psi_L + \psi_R, \tag{2.3}$$

where P_L and P_R are left- and right-handed chiral projection operators and ψ_L and ψ_R are left- and right-handed Weyl fields. Writing the Dirac Lagrangian in terms of the two Weyl fields yield:

$$\mathcal{L}_D = i\bar{\psi_L}\gamma^\mu \partial_\mu \psi_L + i\bar{\psi_R}\gamma^\mu \partial_\mu \psi_R - m(\bar{\psi_R}\psi_L + \bar{\psi_L}\psi_R),$$

which shows that the Dirac Lagrangian can be written using just one Weyl field alone describing a massless Weyl fermion with two degrees of freedom. In the SM, the neutrino is described as a two-component Weyl fermion which violates parity but preserves Lorentz invariance [12]. When introducing a right-handed neutrino field, the Dirac Lagrangian above is splitted into a part in agreement with the SM and a part which is BSM:

$$\mathcal{L}_{\nu\text{-SM}} = i\bar{\psi_L}\gamma^\mu \partial_\mu \psi_L, \tag{2.4}$$

$$\mathcal{L}_{\nu\text{-BSM}} = i\bar{\psi_R}\gamma^\mu \partial_\mu \psi_R - m(\bar{\psi_R}\psi_L + \bar{\psi_L}\psi_R), \tag{2.5}$$

which shows that a right-handed field is necessary in order to generate the Dirac mass term via the Yukawa interaction. The sum of the SM and BSM equations above is identical to the kinetic and potential Lagrangian terms for any of the other massive fermions in the SM when the Dirac spinor is expanded into left- and right-handed components. This is not sufficient to explain why the mass of the SM neutrinos is smaller than 2 eV whereas the mass of the electron is 511 keV as comparison. In order to explain the smallness, it is necessary to introduce Majorana properties for the right-handed neutrino. This is possible since the right-handed neutrinos transform as singlets under the SM gauge transformations so any term including only right-handed neutrinos can be introduced without breaking the gauge invariance of the SM. A Majorana fermion is a fermion which is its own anti-particle: [12]

$$\chi = \chi^c, \tag{2.6}$$

for which χ is used to describe the Majorana field and χ^c is the charge conjugate of the field corresponding to the anti-particle of χ. The Majorana mass term in the Lagrangian can be expressed as:

$$\mathcal{L}_{\nu\text{-BSM}} = -\frac{1}{2} m_R(\bar{\chi^c}\chi + \bar{\chi}\chi^c). \tag{2.7}$$

Now, three extra BSM terms are to be introduced in the SM for the νMSM: a kinematic term for the right-handed neutrinos; a Dirac mass term using the Yukawa interaction of the SM active neutrinos and the BSM sterile neutrinos; and finally a Majorana mass term for the BSM sterile neutrinos. In the simplified case with only one active neutrino denoted ψ_L and one sterile neutrino denoted χ, the BSM Lagrangian density is:

$$\mathcal{L}_{\nu\text{-BSM}} = i\bar{\chi}\gamma^{\mu}\partial_{\mu}\chi - m_D(\bar{\chi}\psi_L + \bar{\psi_L}\chi) - \frac{1}{2}m_R(\bar{\chi^c}\chi + \bar{\chi}\chi^c) \tag{2.8}$$

In the next section, the general expression is examined for which more than one sterile fermion field is introduced. Furthermore, the principles behind the see-saw mechanism will be presented.

2.2.2 The see-saw mechanism

In the previous section, it was shown that a mass term for the neutrino field can be generated by introducing a right-handed neutrino field with Majorana properties. Now, a more general form of the BSM part of the Lagrangian density is studied. When introducing $\mathcal{N}$ right-handed neutrinos, the Lagrangian has the form: [4]

$$\mathcal{L} = \mathcal{L}_{SM} + i\bar{N}_I\partial_{\mu}\gamma^{\mu}N_I - \left(\lambda_{\alpha I}\bar{L}_{\alpha}N_I\tilde{H} - \frac{M_I}{2}\bar{N}_I^c N_I + h.c.\right), \tag{2.9}$$

where the applied variables represent:

N_I	:	Right-handed neutrino singlet field.
I	:	Index over the number of right-handed neutrinos: $I = 1, .., \mathcal{N}$.
L_{α}	:	Left-handed leptonic SU(2) doublet.
α	:	Index used to label the active lepton generation $\alpha = e, \mu, \tau$.
$\lambda_{\alpha I}$	:	Yukawa coupling between the active neutrino in generation α and the I'th right-handed neutrino.
H	:	Higgs field represented by a SU(2) doublet.
M_I	:	Majorana mass for the I'th right-handed neutrino.
$h.c.$	:	Hermitian conjugate of the terms.

Due to the non-zero vacuum expectation value of the Higgs boson, the Dirac masses M_D are defined by the Yukawa coupling and $\langle H \rangle$. In the case where three right-handed neutrinos are added to the SM, a 6x6 mass matrix for neutrinos is present in the Lagrangian. It represents 3 masses for the active SM neutrinos in the three generations e, μ, and τ and 3 masses for the new right-handed neutrinos. The generated Majorana/Dirac mass terms in Eq. 2.9 can be written as: [10]

$$-\frac{1}{2}\begin{pmatrix}\overline{\nu_L} & \overline{\nu_R^c}\end{pmatrix}\begin{pmatrix}0 & \mathbf{M}_D^T \\ \mathbf{M}_D & \mathbf{M}\end{pmatrix}\begin{pmatrix}\nu_L^c \\ \nu_R\end{pmatrix} + h.c., \tag{2.10}$$

where $\mathbf{M}_D$ is a 3×3 matrix of the Dirac masses and $\mathbf{M}$ a 3×3 matrix of the Majorana masses for the right-handed neutrinos. By diagonalizing the mass matrix, the physical state of the system can be found. In the limit where the Majorana masses M_I is much larger than the Dirac masses M_D, the see-saw mechanism gives rise to 3 almost pure right-handed neutrinos with heavy Majorana masses: [5]

$$m_{\nu I}^{\mathrm{RH}} \approx M_I, \tag{2.11}$$

and to 3 almost pure left-handed neutrinos with light Majorana masses

$$\mathbf{m}_{\nu}^{\mathrm{LH}} = -\mathbf{M}_D^T\mathbf{M}^{-1}\mathbf{M}_D. \tag{2.12}$$

The SM neutrino masses are thereby suppressed by the factor $\mathbf{M}^{-1}$ which explains the smallness of the observed neutrino masses. In order to theoretically describe the observed neutrino oscillations, it can be sufficient to introduce two right-handed neutrinos to the SM. But in order to also have a dark matter candidate among the sterile neutrinos, it is necessary to introduce three right-handed neutrinos [4]. Furthermore, if three right-handed neutrinos are introduced, the neutrino sector follows the symmetry seen for the charged leptons and quarks where every left lepton or quark has a right counterpart. When introducing three right-handed neutrinos, 18 free parameters arise: 3 Majorana masses, 3 Dirac masses, 6 mixing angles between sterile and active neutrinos and 6 CP-violating phases [4]. These free parameters are examined in the next section.

2.2.3 The νMSM parameters

The νMSM is defined as the extension of the SM with three sterile Majorana neutrinos represented by right-handed Weyl fields for which the 18 free parameters are tuned to explain all confirmed BSM phenomena. The model is thereby required to e.g. provide a mechanism which can generate the observed baryon asymmetry in the universe and provide a dark matter candidate. The observed neutrino oscillations also set limits for the possible mixing angles between the active neutrinos and sterile neutrinos based on the observed mixing angles for the active neutrinos [10]. All these constraints on the model are examined in [4]. Here it is shown that the νMSM indeed has the potential to explain all the observed BSM phenomena. In this section, I will briefly present some of the restrictions set on the parameters of the model, which are relevant for this analysis, based on the findings in [4].

The three sterile neutrinos are labeled N_1, N_2, and N_3. As mentioned above, the νMSM should provide a dark matter candidate. The lightest sterile neutrino, referred to as N_1, which has an extremely weak coupling to the active neutrinos, is selected as the "dark matter sterile neutrino". The estimated restriction on its lifetime is that it exceeds 10^{24} s which is six orders of magnitude longer than the age of the universe. Due to the extremely weak coupling between the SM neutrinos and N_1, it is not directly relevant for the analysis presented in this thesis which will focus on N_2 and N_3.

The restrictions on the mass of the two additional sterile neutrinos is set by the required baryon asymmetry generation: N_2 and N_3 should have a mass in the range 150 MeV$\lesssim M_{2,3} \lesssim$ 100 GeV and be degenerate: $\Delta M_{2,3} \ll M_{2,3}$. Furthermore, the lifetime of the neutrinos should be $\tau < 0.1$ s. This restriction is set in order to not affect the predictions of big bang nucleosynthesis.

The sterile neutrinos can only be indirectly observed through its coupling to the active neutrinos. Therefore, the potential to discover sterile neutrinos is very dependent on the strength of the coupling between the active neutrinos and the new sterile neutrinos. The mixing is characterized by the coupling strength: $V_{\alpha I}$, where $\alpha = e$, μ, τ again represents the generation of the active neutrino and I the sterile neutrino index. In Ref. [14], the strength of the Yukawa coupling between the active and sterile neutrinos is examined. It is found that the couplings of N_2 and N_3 with the active neutrinos must be almost identical:

$$V_{\alpha 2} \approx V_{\alpha 3}. \tag{2.13}$$

Furthermore, it is found that the coupling $V_{\alpha I}$ highly depends on the mass hierarchy of the active neutrinos which is still unknown. For different scenarios, the sterile neutrino can have a stronger coupling to the e, μ or τ generation of the active neutrinos. In this analysis, the coupling between the sterile neutrinos $N_{2,3}$ and ν_μ is considered. In the next section, the production mechanism and experimental signature for the sterile neutrinos $N_{2,3}$ will be examined. Since the coupling between the two sterile neutrinos N_2 and N_3 and the the active neutrinos ν_α are expected to be of the same size, the coupling strength will be referred to as $V_{\alpha N}$.

3 Sterile neutrino production and cross section

When searching for sterile neutrinos, the search is conducted in a decay channel of the sterile neutrino. In this analysis, the coupling between the sterile neutrino and the SM neutrino ν_μ is investigated. The two sterile neutrinos in the νMSM which were denoted N_2 and N_3 in the previous section are considered. They will also be referred to as Heavy Neutral Leptons (HNL). Due to their probably degenerate mass and coupling strength to the active neutrinos, the two probed sterile neutrinos will be referred to as one and called a HNL or simply N. The coupling strength between the HNL and the active muon neutrino is labeled $V_{\mu N}$ and is used to determine the production rate for the sterile neutrino channel. The coupling strength is also referred to as the mixing angle. In this section, the signal channel which is considered in this analysis is first presented. Thereafter, the cross section for the signal region channel is examined. And finally, the considered mixing angles and sterile neutrino masses are presented.

Figure 1 shows the Feynman diagram of the signal channel used in this analysis. The Hermitian conjugate to this channel is also contributing to the considered SR so the following two final states are included: $W^\pm \to \mu^\pm \mu^\pm e^\mp \nu_e$. The SM neutrino, which will mix into the sterile neutrino, is created through its coupling to an on-shell W-boson which is a gauge boson of the weak force: $W^+ \to \mu^+ \nu_\mu$. The SM neutrino mix with a sterile neutrino which due to its Majorana properties can mix back into the anti-particle to ν_μ with opposite lepton number denoted $\overline{\nu}_\mu$. Thereafter, the SM neutrino decays through a weak interaction: $\overline{\nu}_\mu \to \mu^+ W^{*-} \to \mu^+ e^- \overline{\nu}_e$.

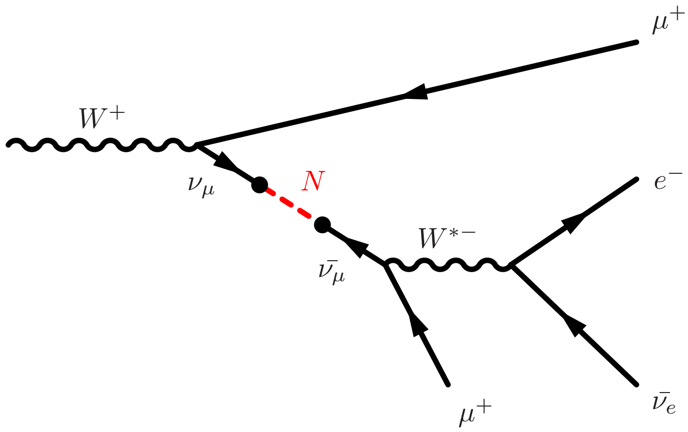


Figure 1: The Feynman diagram for the signal region channel considered in this analysis. Also the Hermitian conjugate of this channel contribute to the signal. The figure is from [15]

The total cross section for the signal region channel is denoted $\sigma(pp \to \mu^\pm \mu^\pm e^\mp \nu_e)$, where pp refers to the two colliding protons. In the following, I will refer to the more general final state channel: $W^\pm \to l^\pm_\alpha l^\pm_\alpha l^\mp_\beta \nu_\beta$ with $\alpha \neq \beta$, for which $\alpha = \mu$ and $\beta = e$ in this analysis. The cross section can be expressed as:

$$\sigma(pp \to l^\pm_\alpha l^\pm_\alpha l^\mp_\beta \nu_\beta) = \sigma(pp \to W^\pm) \cdot \mathrm{Br}(W^\pm \to l^\pm_\alpha N) \cdot \mathrm{Br}(N \to l^\pm_\alpha l^\mp_\beta \nu_\beta), \quad \alpha \neq \beta, \tag{3.1}$$

where the two branching ratios rely on the predictions by the νMSM and are thereby not known. In order to evaluate the branching ratio leading to the BSM sterile neutrino production, the two free parameters describing the HNL mass point, M_N, and the mixing angle, $V_{\alpha N}$, are applied: [16]

$$\mathrm{Br}(W^\pm \to l^\pm_\alpha N) = \mathrm{Br}(W^\pm \to l^\pm_\alpha \nu_\alpha) \cdot |V_{\alpha N}|^2 \cdot \left(1 - \frac{M_N^2}{M_W^2}\right)^2 \left(1 + \frac{1}{2}\frac{M_N^2}{M_W^2}\right). \tag{3.2}$$

The branching ratio for the HNL decay is a function of the partial decay width of the process and the total decay width of the HNL:

$$\mathrm{Br}(N \to l^\pm_\alpha l^\mp_\beta \nu_\beta) = \frac{\Gamma(N \to l^\pm_\alpha l^\mp_\beta \nu_\beta)}{\Gamma}, \tag{3.3}$$

where the total decay width Γ is the sum of all the HNL partial decay widths,

$$\Gamma = \sum_i \Gamma_i(M_N, |V_{\mu N}|^2). \tag{3.4}$$

In order to determine the branching ratio of the HNL decay, it is thereby necessary to calculate all partial decay widths for the HNL in their dependence to M_N and $|V_{\alpha N}|^2$. This is done for the partial decay widths for both the charged-current and neutral-current weak interactions in [17]. Given both the branching ratio for the HNL production and the HNL decay in Eq. 3.2 and 3.3 respectively, the cross section in Eq. 3.1 is determined as a function of the two free parameters M_N and $|V_{\alpha N}|^2$.

The relationship between the three variables: M_N, $|V_{\alpha N}|^2$ and $c\tau$ is now described. The proper decay length $c\tau$ is an important variable for the search of HNLs since it determines the displacement of the HNL vertex from the hard scattering event and thereby has a major impact on how the analysis is performed. The proper decay length is connected to the total decay width for the HNL:

$$c\tau = \frac{c\hbar}{\Gamma}, \tag{3.5}$$

where the total decay decay width can be determined as a function of the variables M_N and $|V_{\alpha N}|^2$ as described above. For given M_N and $c\tau$ one can therefore determine the mixing angle correspondingly. And given M_N and $|V_{\alpha N}|^2$, the cross section can be calculated using the expressions in Eq. 3.1 and 3.2. The total number of predicted signal events is finally given by the product between the cross section, $\sigma(pp \to \mu^\pm \mu^\pm e^\mp \nu_e)$, scaled by the integrated luminosity:

$$N_{\text{events}} = \sigma(pp \to \mu^\pm \mu^\pm e^\mp \nu_e) \int \mathcal{L} \mathrm{d}t. \tag{3.6}$$

In this analysis, the dataset corresponds to an integrated luminosity of 79.8 fb^{-1}.

3.1 Signal samples

In this analysis, a prompt trilepton search for HNL signals is conducted. Five signal samples are used to investigate possible sterile neutrino signatures in the ATLAS detector. The assumption that the signal signature is prompt requires that the decay length of the HNL is small. A dedicated search for HNLs with small mass and long decay length is conducted by another analysis group within the ATLAS Collaboration which uses displaced vertices to probe the signal [18]. The properties of the applied samples are first presented.

The considered HNL masses are $M_N = 5$, 10, 20, 30 and 50 GeV. In the previous section, it was described how the proper decay length, $c\tau$, of the HNL is related to the HNL mass and mixing angle. This relationships is visualized in Fig. 2 which shows the proper decay length contours for the HNL as a function of $|V_{\mu N}|^2$ and M_N. The black line shows the limit set by the DELPHI experiment in the region [19]. Since this figure was made, the CMS experiment has set new limits in this region based on the 2015-2016 Run 2 data of the LHC [22] which is just below the DELPHI limit. The plot shows that the proper decay length decreases when the mixing angle increases.

The signal parameters are chosen based on two considerations: a) $c\tau$ needs to be small for the signal to be prompt, and b) the mixing angle should be smaller than the current limits for the HNL sensitivity to be able to improve the knowledge of HNLs. This is easiest achieved for the larger mass points which are more unstable and therefore tend to have a smaller decay length. The black boxes in Fig. 2 shows the parameters used in the five generated MC samples.

The mixing angles corresponding to the considered HNL mass points M_N and proper decay lengths $c\tau$ are determined as described in the previous section. Table 3 shows the variables used for the five signal samples. For each of the listed HNL mass points M_N (with corresponding mixing angle

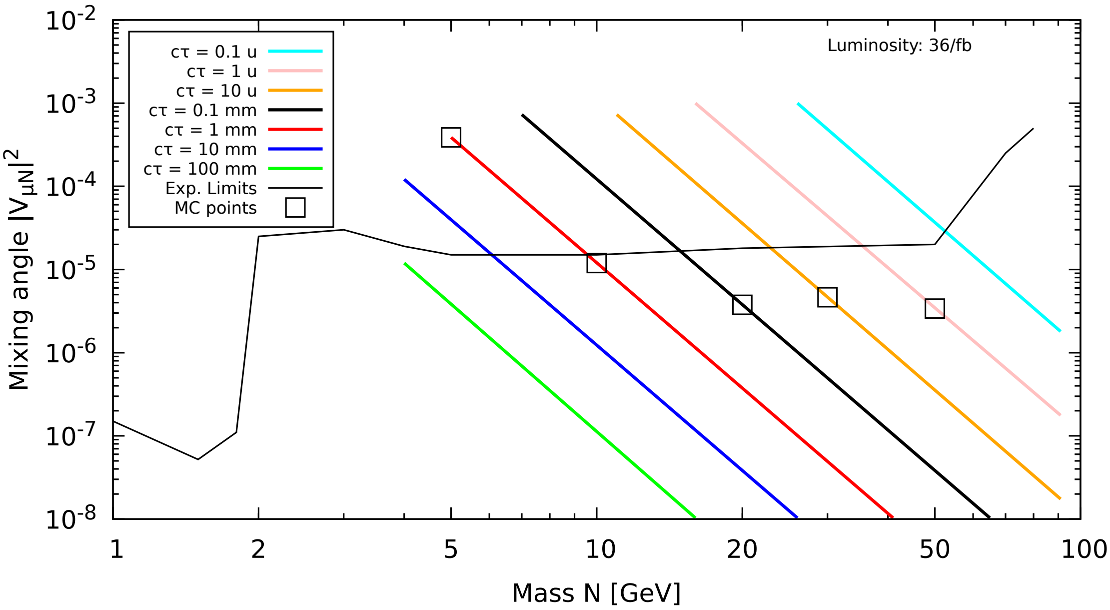


Figure 2: The proper decay length contours for the HNL as a function of the mixing angle, $|V_{\mu N}|^2$, and the mass of the HNL, M_N. The five boxes indicate the applied values of M_N, $|V_{\mu N}|^2$ and $c\tau$ used for the generated MC signal samples. The black line indicates the exclusion limit in this region which is set by the DELPHI experiment [19] and the ATLAS [20] and CMS [21] experiments at $\sqrt{s} = 8$ TeV. Since this figure was made, the CMS experiment has set new limits in this region based on the 2015-2016 Run 2 data of the LHC which lies just below the limit shown [22]. The phase space above the black line has thus already been experimentally excluded. The figure is from [15]

Chosen parameters for the five MC signal samples						
M_N	:	5 GeV	10 GeV	20 GeV	30 GeV	50 GeV
$c\tau$	:	1 mm	1 mm	0.1 mm	10 μm	1 μm
$\lvert V_{\mu N}\rvert^2$	:	$3.9 \cdot 10^{-4}$	$1.2 \cdot 10^{-5}$	$3.8 \cdot 10^{-6}$	$4.6 \cdot 10^{-6}$	$3.4 \cdot 10^{-6}$

Table 3: The parameters used to generate the five MC signal samples, where M_N is the mass of the HNL, $c\tau$ is the proper decay length, and $|V_{\mu N}|^2$ is the mixing angle between the SM neutrino ν_μ and the HNL.

$|V_{\mu N}|^2$), the cross section of the signal can be determined using Eq. 3.1.

When the mixing angle is increasing, the cross section of the signal signature is increasing and thus the number of predicted signal events is increasing. Therefore, a higher luminosity will in general enable the analysis to exploit lower mixing angles. Though for a constant HNL mass point, the proper decay length increases when the mixing angle decreases which will change the efficiency of the analysis.

4 The LHC and the ATLAS detector

In this section, the layout and main properties of the Large Hadron Collider (LHC) at CERN and the ATLAS (A Toroidal LHC ApparatuS) detector will be introduced. First, the LHC is shortly described giving a brief overview of its design and performance. The complete description of the LHC can be found in the LHC Design Report [23–25]. Secondly, the key features of the ATLAS detector are presented. A detailed description of the ATLAS experiment can be found in [26]

4.1 The LHC

The LHC [23–25] is the largest and most powerful particle accelerator in the world which is run by the European Organization for Nuclear Research (CERN). CERN was founded in 1954 as an international research organization which now has 22 member states and over 600 institutions and universities world wide use the CERN facilities [7]. The LHC is the largest accelerator run by CERN and the majority of the CERN activities involve the experiments around the LHC. The aim of the LHC is to get a better understanding of the elementary particles and their interactions.

The LHC lies close to Geneva and it crosses the border between Switzerland and France. It has a circumference of 27 km and is placed in a tunnel 100 meters below ground level. It is a two-ring, superconducting accelerator which was build to run with a centre-of-mass energy of up to 14 TeV. Both protons and heavy ions can be injected in the LHC. This analysis is based on proton-proton collisions from the LHC so focus will be on these events. The four large detectors at the LHC are ATLAS, CMS, ALICE and LHCb, which are located at the LHC beam intersection points. Both ATLAS and CMS are general-purpose detectors designed to study the properties of the SM, especially the Higgs boson, and possible signs of new physics.

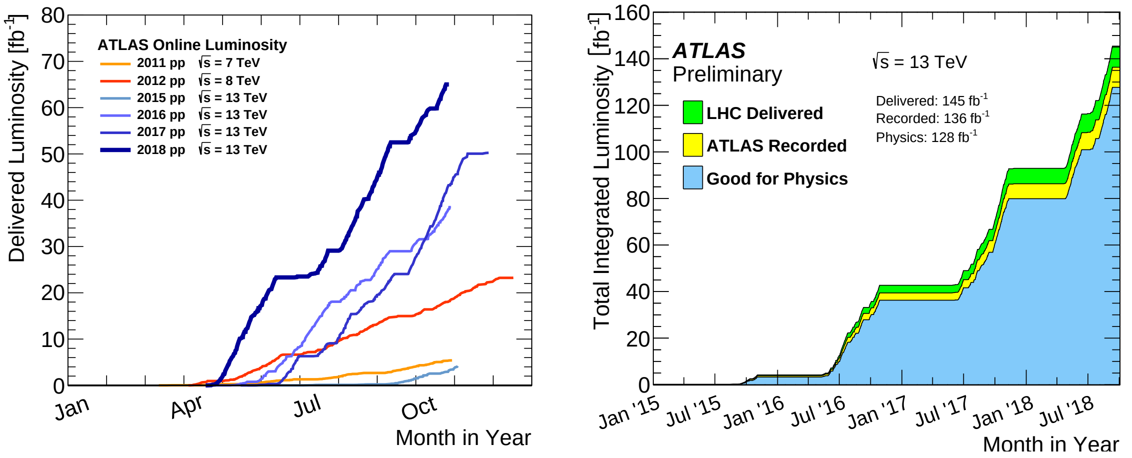
The two most important features of a particle accelerator are the centre-of-mass energy, $\sqrt{s}$, and the luminosity, $\mathcal{L}$, which determine the interactions which can be studied and the event rates. For a given process, the number of interactions N is given by the product between the total luminosity and the cross-section for the process σ:

$$N = \sigma \int \mathcal{L} \mathrm{d}t. \tag{4.1}$$

The integrated luminosity, $\int \mathcal{L}\mathrm{d}t$, will for simplicity be referred to as $\mathcal{L}$.

The experiments at the LHC started taking data in December 2009 which was the beginning of Run 1 lasting until the end of 2012. During Run 1, the LHC increased the centre-of-mass energy from 2.36 TeV to 8 TeV. Following Run 1, the LHC shut down for two years in 2013-2014. Doing the Long Shutdown 1, all the detectors were upgraded, the LHC was extensively renovated, and the computing resources were scaled to meet higher luminosity challenges [7]. In 2015, the Run 2 of the LHC started which will last until the end of 2018. Doing Run 2, the LHC has been operating at a centre-of-mass energy of $\sqrt{s} = 13$ TeV. It has successfully been running in the period from 2015-2018 and delivered stable results at high energies and with an increasing luminosity. Figure 3a shows the integrated luminosity by the LHC for both the Run 1 data-taking years 2011 and 2012 and for the Run 2 years 2015-2018. Figure 3b similarly shows both the total luminosity for Run 2 and the total luminosity recorded by the ATLAS detector. It is found that 94% of the integrated luminosity was recorded by the ATLAS experiment and a total of 88% is good quality data used for analyses. In this analysis, the Run 2 dataset recorded by the ATLAS experiment from 2015-2017 is applied which corresponds to an integrated luminosity of 79.8 fb^{-1}.

The high luminosity delivered by the LHC and its resulting high interaction rate is crucial for analyses of particle processes with small cross-sections. It gives the best conditions ever seen to measure rare interactions and search for potential new physics as presented in this analysis. The high luminosity also puts a lot of pressure on the detectors both in order to record and process the large amount of data and for the ability to accurately measure the properties of the particles associated to each event. Figure 4 shows the integrated luminosity as a function of the mean number of interactions per bunch crossing denoted $\langle\mu\rangle$. It is shown that $\langle\mu\rangle$ =13.4, 25.1 and 37.8 for the data taken in 2015, 2016 and 2017 respectively. The impact of additional interactions in the same or neighboring bunch crossings is referred to as pile-up [28]. In proton-proton collisions, the cross-section for QCD jet production processes is heavily dominating compared to the cross sections of the more rare physics processes being searched for in the particle experiments. So a highly efficient triggering system is crucial for the detectors in order to select which events should be kept. In the next section, I will examine some of the properties of the ATLAS detector among which is the trigger system.



(a) The integrated luminosity for each data taking year from 2011-2012 in Run 1 and 2015-2018 in Run 2 as a function of the month in the year. The delivered luminosity by the LHC is measured in fb^{-1}.

(b) The integrated luminosity for Run 2 of the LHC as a function of time from 2015-2018. In addition to the total luminosity delivered by the LHC, the luminosity recorded by the ATLAS experiment is shown.

Figure 3: Integrated luminosity delivered by the LHC. The figures are published by the ATLAS Collaboration in [27].

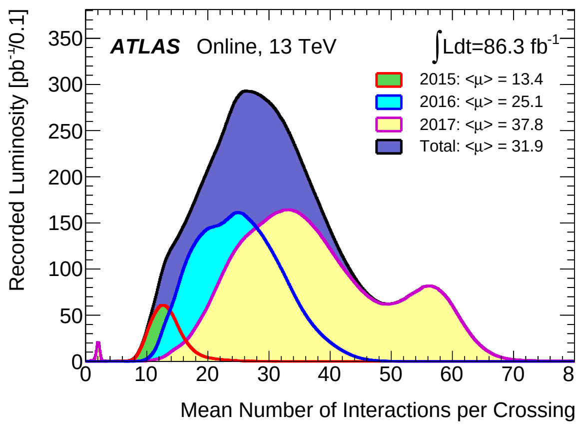


Figure 4: The recorded luminosity as a function of the mean number of interactions per crossing measured by the ATLAS detector. The figure is published by the ATLAS Collaboration in [27].

4.2 The ATLAS detector

ATLAS is a particle detector constructed at the LHC [26]. In this section, the structure of the ATLAS detector is first presented including both the coordinate system applied when referring to the detector and the four main sub-detectors. Thereafter, the trigger system of the ATLAS experiment is introduced. Finally, the experimental signature of particles in the detector is briefly discussed.

The ATLAS detector has a cylindrical geometry and is forward-backward symmetric. Figure 5 shows a schematic view of the detector and its main components. In the middle of the cylindrical detector is the beam line which is referred to as the z-axis in the ATLAS coordinate system. The origin of the coordinate system is defined as the interaction point where the two proton bunches meet in the center of the detector. The x-y plane is transverse to the beam direction for which the positive x-axis is defined as pointing from the interaction point to the centre of the LHC ring and the positive y-axis is pointing to the sky. The azimuthal angle ϕ is measured around the beam axis, the polar angle θ is the angle from the beam axis, and finally is the pseudorapidity defined as $\eta = -\ln\tan(\theta/2)$. The four main detector components which are shown in Fig. 5 are: the inner tracking detector (ID) at the core surrounded by a thin superconducting solenoid generating

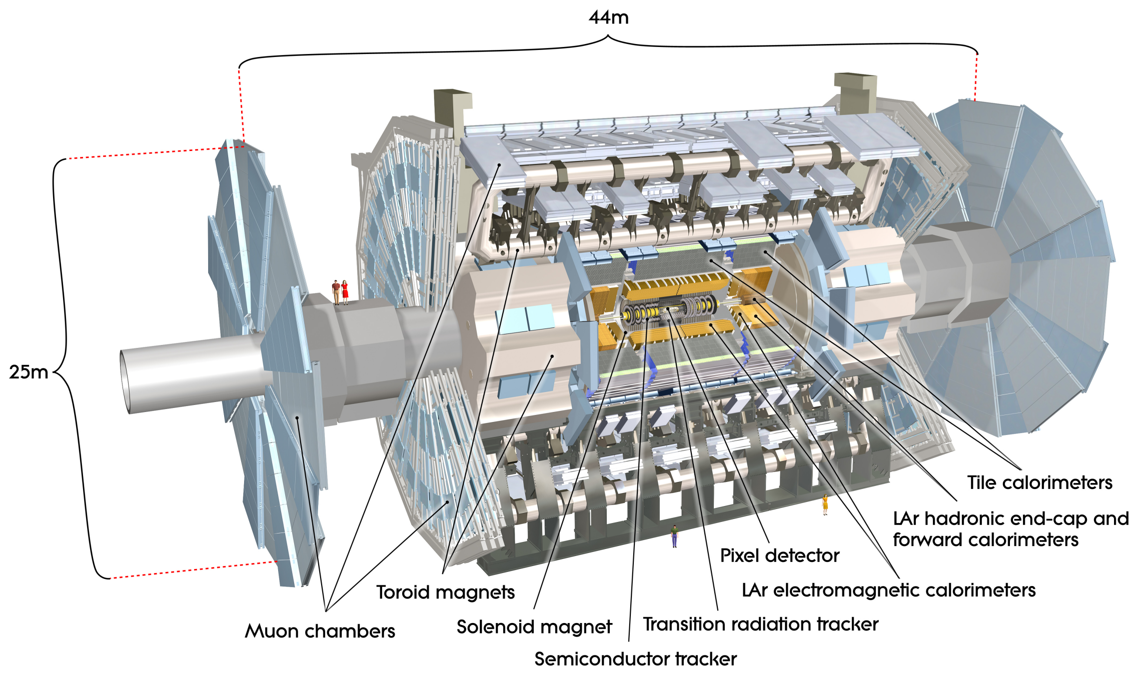


Figure 5: Schematic view of the ATLAS detector where the main detector components are labeled. Figure from the ATLAS experiment official website in [29].

a 2 T axial magnetic field; the electromagnetic and hadronic calorimeters; and finally the muon spectrometer incorporating three large superconducting toroid magnets. These four sub-detectors are introduced below. A detailed overview of the detector details can be found in [26].

The ID consists of three sub-detectors which provide charged-particle tracking and momentum measurement in the pseudorapidity range $|\eta| < 2.5$. A silicon pixel detector supplies high-granularity measurements of the track, including the insertable B-layer [30, 31] installed after Run 1 of the LHC, which is followed by a silicon microstrip tracker also referred to as a semiconductor tracker. At larger radii, a transition radiation tracker enables radially extended track reconstruction up to $|\eta| = 2.0$, which enhances the pattern recognition and momentum resolution, and provides electron identification information over a wide range of energies.

The calorimeter system covers the pseudorapidity range $|\eta| < 4.9$. Electromagnetic (EM) calorimetry is provided by high-granularity lead/liquid-argon (LAr) detectors in the pseudorapidity range $|\eta| < 3.2$, where it is divided into a barrel part ($|\eta| < 1.475$) and two end-cap components ($1.375 < |\eta| < 3.2$). The transition region between the two EM calorimeters is known as the crack region. In this region, a significant quantity of material is contained which leads to energy loss and a reduction in the performance. Therefore, the electrons in the analysis are required to be outside the crack region. The EM calorimeter is extended to larger pseudorapidities by the forward calorimeter which consists of an EM and hadronic part. The EM and hadronic calorimeter parts of the forward calorimeter are copper-LAr and tungsten-LAr detectors respectively. Hadronic calorimetry is provided by the steel/scintillator-tiles calorimeter in the region $|\eta| < 1.7$, which is segmented into three barrels: a central barrel and two extended barrels. It is extended up to $|\eta| < 4.9$ by the forward calorimeter as mentioned above.

The muon spectrometer includes separate trigger and high-precision chambers. The magnetic field for the muon spectrometer, which deflects the muon tracks, is generated by three large air-core toroids. The chamber system covers the pseudorapidity region $|\eta| < 2.7$, where precision measurements of the track coordinates are provided by monitored drift tubes, and at large pseudorapidities $2 < |\eta| < 2.7$, the cathode strip chambers with higher granularity are used. The muon trigger system covers the range $|\eta| < 2.4$, where resistive plate chambers are used in the barrel and thin gap chambers in the end-cap regions.

A two-level trigger system is used to select events of interest [32]. The data from the detector-specific readout electronics is received by the data acquisition system. A subset of the detector information is used by the first Level-1 trigger. The Level-1 (L1) trigger is implemented in hardware and uses the detector information to make a quick decision of whether the event will be kept. The L1 reduces the event rate from the design bunch-crossing rate to an event rate of around 100 kHz. After the acceptance of the L1 trigger, the events are processed by the software-based high level trigger (HLT) with an output rate of about 1 kHz. The HLT applied in this analysis is described in Sec. 7.1.

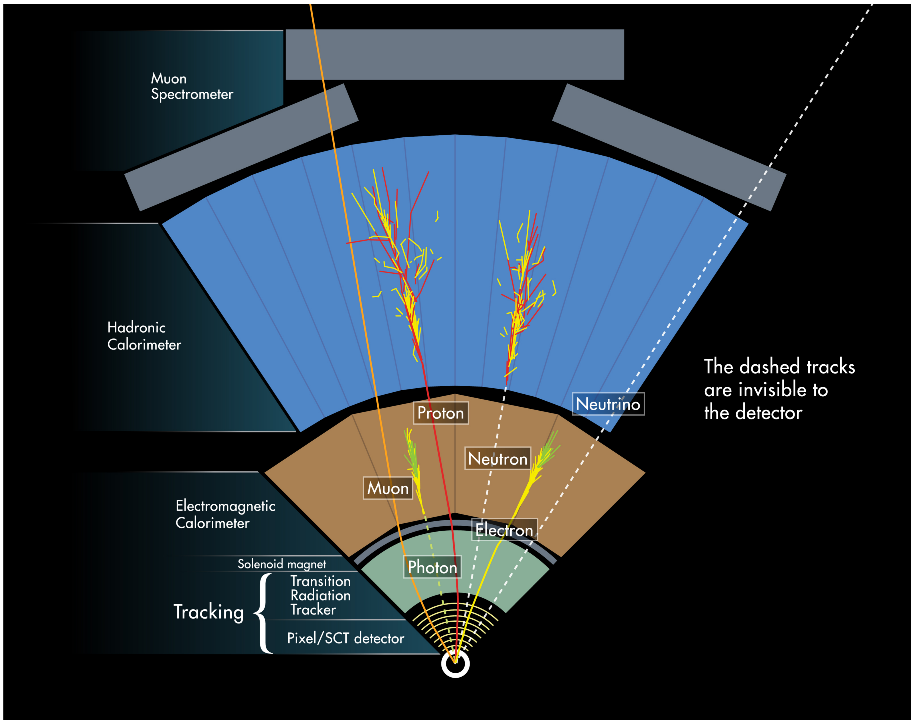


Figure 6: Transverse view of the ATLAS sub-detectors: the inner detector; the electromagnetic calorimeter; the hadronic calorimeter; and the muon spectrometer. The experimental signature of specific particles are shown. A proton is used to illustrate the signature of a charged hadron whereas a neutron is illustrating the signature of a neutrally charged hadron. Hadrons are measured as jets due to the hadronization of quarks and gluons. Photo from the ATLAS experiment official website in [33].

Figure 6 shows the experimental signature of particles in the four main parts of the sub-detectors. It is illustrated that only charged particle tracks are measured in the ID which in this figure are the muon, proton and electron. Electrons and photons are measured in the EM calorimeter whereas jets are measured in the hadronic calorimeter. A proton is used to illustrate the signature of a charged hadron and a neutron as a neutrally charged hadron. Due to the hadronization of quarks and gluons, they are measured as jets in the detector which are reconstructed from the calorimeter cell signals using a clustering algorithm [34]. Muons pass both the EM and hadronic calorimeters and are measured in the muon spectrometer. Neutrinos are the only known SM particle not measured at all in the detector. By calculating the missing transverse energy in each event, the unmeasured transverse energy from neutrinos, and possibly other BSM particles, can be estimated, see Sec. 6.1.4.

5 Data and Simulation Samples

In this section, the data and MC samples are described. The dataset recorded by the ATLAS experiment is first presented. Thereafter, the main properties of the MC samples are listed.

The data used in this analysis is collected by the ATLAS detector in the period from 2015-2017 and corresponds to an integrated luminosity of $\mathcal{L} = 79.8$ fb^{-1}. The uncertainty in the combined 2015-2017 integrated luminosity is 2.0% [35]. It is derived following a methodology similar to that detailed in Ref. [36]. The data is a part of the Run 2 of the LHC for which the centre-of-mass energy is $\sqrt{s} = 13$ TeV. Table 4 shows the integrated luminosity and average number of interactions per bunch crossing for the data-taking periods applied in the analysis. When producing the samples for the analysis, Good Run Lists [37] are applied which ensures that only events which are of good quality for physics analyses are applied. The integrated luminosity used in this analysis is calculated based on the Good Run Lists.

Year	$\int \mathcal{L}\mathrm{d}t$ [fb^{-1}]	Average μ
2015	3.2	13.4
2016	33.0	25.1
2017	43.6	37.8

Table 4: The data-taking periods applied in this analysis. Both the integrated luminosity [37] labeled $\int \mathcal{L}\mathrm{d}t$ and the average number of interactions per bunch crossing [27] labeled "average μ" are shown for each year.

MC simulations are used in the analysis in order to model both the known SM background processes based on the SM theory and the signal samples based on the νMSM. The MC events are simulations of events produced in the LHC including the detector response which is modeled using the ATLAS detector simulation based on GEANT4 [38]. Four MC generators are applied in the study: PYTHIA [39], POWHEG [40], MADGRAPH [41] and SHERPA [42, 43]. The MC samples are designed to match the data conditions. This includes information about the pile-up conditions and the triggers used for the data. The applied MC samples were launched after the data taking was complete so the measured conditions from data are applied to get an accurate description of data.

In Tab. 5, all the applied MC samples are listed together with both the MC generator used to produce the samples, the set of Parton Distribution Functions (PDF) used, and a reference to detailed tables in appendix Sec. A for each sample. The PDFs are used to define the structure of the protons in the pp collisions. They provide the probability densities for each of the possible partons in the hadron as a function of the momentum fraction and energy scale [44].

Process	MC generator	PDF	Appendix
Signal	Pythia v8.2	NNPDF2.3LO	Tab. 32-33
$t\bar{t}$	Powheg + Pythia v8.2	NNPDF2.3LO	Tab. 34-35
$t\bar{t}W$	MadGraph5 v2.3 + Pythia v8.2	NNPDF2.3LO	Tab. 36-37
$t\bar{t}Z$	MadGraph5 v2.2 + Pythia v8.1	NNPDF2.3LO	Tab. 36-37
t and Wt	Powheg + Pythia v6.4	Perugia2012	Tab. 38-41
W+jets	Sherpa v2.2	NNPDF3.0NLO	Tab. 42-47
Z+jets	Sherpa v2.2	NNPDF3.0NLO	Tab. 48-53
$V\gamma^*$	Sherpa v2.2	NNPDF3.0NLO	Tab. 54-55
VV: fully leptonic	Powheg + Pythia v8.2	AZNLO CTEQ6L1	Tab. 56-61
VV: partial hadronic	Sherpa v2.2	NNPDF3.0NLO	Tab. 56-61
VVV: fully leptonic	Sherpa v2.2	NNPDF3.0NLO	Tab. 62-63
VVV: partial hadronic	Powheg + Pythia v8.1	AZNLO CTEG6L1	Tab. 62-63

Table 5: List of all the MC processes applied in the analysis and the Monte Carlo (MC) generators and parton distribution functions (PDF) applied for the processes. Details of each MC sample is shown in appendix Sec. A. MC processes with a V in the process description refers to a W or Z-boson. The diboson samples denoted VV are: WZ, WW, and ZZ. The triboson samples denoted VVV are: WWW and ZWW.

6 Basic concepts for defining the analysis procedure

When two protons collide in the ATLAS detector, new particles are created. The particles traverse the detector and leave signals in the layers of the detector which outline the trajectory of the particles. In the previous section, the sub-detectors of the ATLAS detector were briefly examined. Here it was discussed which parts of the detector the different particles leave signals in. Objects are reconstructed and identified from the detector signals using algorithms. In this section, the reconstruction and identification algorithms for muons, electrons, and jets are presented. These algorithms are defining the objects used and they are optimized to avoid mis-identification while maintaining a high efficiency.

The fully reconstructed final state of a proton-proton collision consists of the identified individual particles including electrons, photons, muons, tau-leptons, jets, and missing transverse momentum (E_T^{miss}). In this analysis, the final state is required to be two muons and an electron. Both photons, tau-leptons, and jets can also contribute to the number of final state leptons: energetic photons which decay into an electron-position pair $\gamma \to e\bar{e}$; tau-leptons which decay leptonically into an electron or muon, $\tau \to e\bar{\nu}_e\nu_\tau/\mu\bar{\nu}_\mu\nu_\tau$; and jets containing heavy flavour quarks which decay semi-leptonically, e.g. $b \to e\bar{\nu}_e c/\mu\bar{\nu}_\mu c$. In data events, the kinematic properties of the reconstructed electrons and muons is used to indicate the origin of the leptons and whether they are signal-like or background-like. For MC samples, it is additionally possible to match the individual reconstructed particles to their original type and mother particle.

In this section, the object definitions for electrons, muons, jets and missing transverse energy are first presented and the overlap removal between reconstructed electrons, muons and jets described. The isolation definitions and the impact parameter variables are also presented since they will be applied in the signal region requirements for the reconstructed electrons and muons. Thereafter, the MC weights and scale factors used in the analysis are presented. These are the weights applied for each MC events in order to ensure that the MC simulations accurately reflect the corresponding data events. Finally, lepton truth matching is used in MC events to distinguish events with real/prompt leptons from events with fake/non-prompt lepton.

6.1 Object definitions

Object definitions are based on recommendations by Combined Performance groups within the ATLAS Collaboration. For both muons and electrons, the object definitions make requirements to the kinematic properties of the track, the identification and isolation variables, and the impact parameters. In this section, the definitions of the identification and isolation variables and the impact parameters are presented. Thereafter, the specific choices used for each type of object are listed.

6.1.1 Electron reconstruction and identification

Electrons which traverse the detector leave signals in both the inner detector and the EM calorimeter. By matching the tracks formed in the inner detector with the energy clusters in the EM calorimeter, electrons can be reconstructed. The electron identification (ID) algorithm used is the likelihood-based (LLH) method [45]. It is based upon a multivariate analysis technique which consider several properties of the electron candidate simultaneously when determining the overall probability for the candidate to be an electron or not. Different selection levels are defined using the algorithm which vary in how tight the selection of the electrons is by changing the required efficiency. In this analysis, the `LooseAndBlayerLLH` selection is used which require a signal efficiency of 90% for electron candidates with $E_T \approx 25$ GeV [45, 46]. The electron candidate is also required to have a hit in the b-layer of the detector [47] in order to fulfill the selection. Additional requirements are imposed on the transverse momentum and pseudorapidity of the electron candidate which is required to be $p_T > 5$ GeV and $|\eta| < 2.47$ with a veto for the calorimeter crack region for which $1.37 < |\eta| < 1.52$.

6.1.2 Muon reconstruction and identification

Muon tracks are reconstructed based on signals from both hit points in the inner detector and the muon spectrometer and is also using additional information about energy deposits in the calorimeters. When the muon candidate has tracks both in the inner detector and the muon spectrometer, it is denoted `combined` [48]. In this analysis, all muons are required to be `combined`. The identification criteria for the muon can be `loose`, `medium` or `tight` where the names indicate the efficiency requirement for each [48, 49]. For the `loose` selection, the reconstruction efficiency is very high where for the `tight` selection the purity of the muons is larger at the cost of the efficiency. The three sets of selection criteria are inclusive, so muons which satisfy the `tight` selection also satisfy the `medium` and `loose` selection. In this analysis, the `loose` selection is used for the subleading muon in the final state and the `tight` selection is used for the leading muon. Further requirements for the transverse momentum and pseudorapidity of muon candidates are $p_T > 5$ GeV are $|\eta| < 2.5$.

6.1.3 Jet reconstruction and identification

Jets traversing the detector are characterized using their bundles of tracks in the inner detector and the locally-calibrated topological clusters which are built from calorimeter cells. Jet reconstruction is performed from the topological clusters in the calorimeters using the anti-k_T algorithm with radius parameter $\Delta R = 0.4$ ($\Delta R = \sqrt{\Delta\eta^2 + \Delta\phi^2}$) [50, 51]. The requirements made for the transverse momentum and pseudorapidity of jets are: $p_T > 20$ GeV and $|\eta| < 4.5$.

Jets from pile-up are suppressed using a selection based on the jet vertex tagger (JVT) algorithm [52]. The JVT algorithm is a tagging algorithm for jets used to identify jets coming from the hard scattering interaction and thereby reject jets from pile-up. Information about the tracks associated to the jet and information about the vertices are is in the algorithm. The applied JVT selection corresponds to a 92% efficiency in tagging jets and is allowing an observed fake rate of 2% from pile-up jets.

Jets which originate from a b-quark have in general different properties than jets initiated by the lighter quarks so an algorithm is used to tag the b-jets. The algorithm uses a multi-vary approach to tag the relatively long-lived b-jets [53]. In this analysis, the b-jet selection which tags jets with a 77% efficiency is applied.

6.1.4 Missing transverse energy

The missing transverse momentum $\mathbf{p}_T^{\mathrm{miss}}$ with magnitude E_T^{miss} is defined as the negative vector sum of transverse momenta of all identified physics objects (electrons, photons, muons, taus, jets) and an additional soft term [54]. The soft term includes all tracks which are not associated to a physics object but still associated to the primary vertex. In order to get an accurate estimate for the missing energy in the event, it is crucial to have a good pile-up suppression as described above. For many physics analyses, E_T^{miss} is an essential variable since it is used to estimate the missing measurements from SM neutrinos, and some beyond the SM particles, which escape the detector without producing a signal.

In this analysis, there is missing energy in the signal process final state from one SM neutrino. The neutrino is arising from the HNL decay through the leptonic decay of the W-boson so the missing energy in the signal process is correlated with the mass of the HNL. For an increasing invariant mass M_N, the E_T^{miss} tends to increase as well. Even though missing energy is present in the signal sample, the size of the missing transverse energy is much smaller than the typical amount of E_T^{miss} in background processes as $t\bar{t}$. So by setting an upper cut on E_T^{miss}, a reduction in the number of background events can be achieved.

6.1.5 Overlap removal

Overlap removal is used in order to avoid double counting objects when the same particle is reconstructed as different object types. Table 6 shows how the overlap removal is performed for muons, electrons, and jets with respect to each other. Overlap removal between a jet and a lepton is performed by considering the angular separation $\Delta R = \sqrt{\Delta\eta^2 + \Delta\phi^2}$ between the objects and their transverse momentum. Overlap removal between two reconstructed electron and muon objects is performed by considering the tracks and thereby not based on the ΔR separation.

	Overlap removal: electrons and jets	
1.	Jet removed	Jets within $\Delta R < 0.2$ of an electron track.
		Unless it is a b-jet or has 20% higher p_T than the electron.
2.	Electron removed	Electrons within the jet sliding cone defined by
		$\Delta R < 0.04 + 10\ \mathrm{GeV}/p_T(e) < 0.4$.
	Overlap removal: muons and jets	
1.	Jet removed	Same requirement as for electrons.
		Unless it is a b-jet or the jet has three or more tracks.
2.	Muon removed	Same requirement as for electrons.
	Overlap removal: muons and electrons	
1.	Muon removed	When a calorimeter tagged muon shares the same tracks
		as an electron.
2.	Electron removed	If it shares the same track with a remaining muon.

Table 6: Overlap removal requirements between reconstructed muons, electrons and jets with respect to each other. No overlap removal is performed with τ objects or between b-jets and jets.

6.1.6 Isolation

Isolation criteria are applied in order to distinguish leptons originating from a jet, which are likely to have many other tracks in their close proximity, from signal-like leptons. The isolation of a track is defined using a spatial cone around the track with radius $\Delta R = \sqrt{\Delta\eta^2 + \Delta\phi^2}$ and then the p_T of other tracks or calorimeter deposits E_T within the cone are considered [48].

Isolation criteria are one of the main discriminating features between signal-like and background-like leptons used in this analysis. It is also crucial for the fake lepton estimation in Sec. 8 where the discriminating power of different isolation selection criteria is exploited. Therefore, a more detailed presentation of the isolation criteria is given below.

The isolation selection apply cuts on the following variables:

- `topoetcone20/pt` defined as the sum of E_T in the topologically connected clusters in a cone of $\Delta R = 0.2$ around the particle track divided by the transverse momentum of the particle,
- `ptvarcone20/pt` or `ptvarcone30/pt` defined as the sum of track-p_T in a cone of $\Delta R = 0.2$ or $\Delta R = 0.3$ around the particle track divided by the transverse momentum of the particle.

Two different types of selections are used: a "fixed cut" type where the criteria consist of simple cuts on the variables, for example `topoetcone20/pt`< 0.06 and `ptvarcone20/pt`< 0.06; and a targeted efficiency type based on cut maps parametrized in η and p_T of the track [55]. In Tab. 7, the definitions of the different isolation selections are listed. The first four isolation selections are efficiency targeted whereas the additional selections are the "fixed cut" type.

Isolation name	Object	Calorimeter isolation	Track isolation
`LooseTrackOnly`	$e + \mu$	–	$\epsilon = 99\%$
`Loose`	$e + \mu$	$\epsilon = 99\%$	$\epsilon = 99\%$
`GradientLoose`	$e + \mu$	ϵ =(0.057*pT[GeV]+95.57)%	ϵ =(0.057*pT[GeV]+95.57)%
`Gradient`	$e + \mu$	ϵ =(0.1143*pT[GeV]+92.14)%	ϵ =(0.1143*pT[GeV]+92.14)%
`FixedCutLoose`	μ	Cut: topoetcone20/pT<0.3	Cut: ptvarcone30/pT<0.15
`FixedCutLoose`	e	Cut: topoetcone20/pT<0.2	Cut: ptvarcone20/pT<0.15
`FixedCutTightTrack`	μ	–	Cut: ptvarcone30/pT<0.06
`FixedCutTightTrack`	e	–	Cut: ptvarcone20/pT<0.06
`FixedCutTight`	μ	Cut: topoetcone20/pT<0.06	Cut: ptvarcone30/pT<0.06
`FixedCutTight`	e	Cut: topoetcone20/pT<0.06	Cut: ptvarcone20/pT<0.06

Table 7: Definitions of the isolation selections [55]. The combined isolation efficiency for the `LooseTrackOnly` and `Loose` selection is 99%. For the `GradientLoose` and `Gradient` selections, the combined isolation efficiency is $\epsilon = 90\%/99\%$ and $\epsilon = 95\%/99\%$ respectively for $p_T = 25/60$ GeV.

For this analysis, the fixed cut selections called `FixedCutTightTrackOnly` and `FixedCutLoose` are used for the leading and subleading muon respectively, and the efficiency targeted working point `Gradient` is used for the electron. The `Loose` isolation is also applied in the analysis for both the subleading muon and electron in order to increase the statistics and test the stability of the background estimation method in Sec. 8.

6.1.7 Impact parameter

The transverse impact parameter (IP) d_0 and the longitudinal IP z_0 are defined with respect to the beam-line [56]. They represent the shortest distance between the reconstructed particle track and the beam-line which can be recalculated to get the distance between the reconstructed particles track and the primary vertex of the hard scattering event. Variables based on the IPs are used to distinguish particle trajectories that are very likely to origin from the primary vertex with particles that are likely to origin from a decay of particles with a long lifetime or from pile-up. The two variables which are applied in the analysis are defined as:

- $|d_0|/\sigma(d_0)$ which is the smallest transverse distance between the particle trajectory and the beam-line divided by the uncertainty of the distance,
- $|\Delta z_0 \sin(\theta)|$ where $|\Delta z_0|$ is the distance between the primary vertex of the hard scattering event and the particle trajectory, and $\sin\theta$ is included to avoid rejecting tracks with an expected larger error in the forward region of the detector.

For the transverse IP, no transformation is needed since a transverse distance is independent on the position along the beam-line. The corrections to be applied in MC to leptons are derived under the assumption that recommended cuts are made on the two IP variables. The recommended requirements are:

- Muons: $|d_0|/\sigma(d_0) < 3$ and $|\Delta z_0 \sin(\theta)| < 0.5$ mm.
- Electrons: $|d_0|/\sigma(d_0) < 5$ and $|\Delta z_0 \sin(\theta)| < 0.5$ mm.

In Sec. 7.3, it is studied how these IP requirements effects the significance of the signal events. It is found that the recommended cut on $|d_0|/\sigma(d_0)$ introduce an inefficiency for signal events, especially those with the longest proper decay length of $c\tau = 1$ mm. The long decay length affects the origin of the particle trajectories which become displaced from the primary vertex. In order to avoid this inefficiency for signal events, the subleading muon and electron are only required to satisfy $|\Delta z_0 \sin(\theta)| < 1.0$ mm as a preselection requirement. In Sec. 7.3, an optimization study is performed for the SR for which it is found that the optimal requirements are $|d_0|/\sigma(d_0) < 10$ and $|\Delta z_0 \sin(\theta)| < 0.5$ mm for both the subleading muon and the electron.

6.2 MC weights and scale factors

MC events are simulated particle collision events which are based on theoretical predictions. In order to ensure that the MC samples are reflecting data, different weights to the MC events are assigned. First, the MC events are normalized to the correct cross section and scaled to the total integrated luminosity of the data which is 79.8 fb^{-1} in this analysis. Monte Carlo weights are applied which are weights assigned by the generator in order to obtain the kinematic distributions seen in data. Furthermore, pile-up weights are used aiming to imitate the pile-up conditions seen in data. The pile-up weight is an additional weight used to correctly scale the MC events to the pile-up conditions of data.

Finally, scale factors (SFs) are applied in order to increase the accuracy of the MC modeling for the specific region of phase space for the particles considered. SFs take into account the type of particles required in the region and the kinematic properties of these particles. They are derived in bins of transverse momentum and rapidity for the electron/muon and calculated by taking the ratio of the determined efficiency for data and MC: SF = ϵ(data)/ϵ(MC). Four SFs are available for electrons [57] which model the MC efficiency given different lepton requirements: reconstruction; identification; isolation; and trigger. The full electron SF is determined by multiplying the individual SFs. For muons, three SFs are available denoted: reconstruction; isolation; and trigger SF [58]. The reconstruction SF for muons take into account the identification quality selection applied. Two additional SFs are available for jets: one efficiency SF for the applied JVT selection; and one for the b-jet selection. In this analysis, scale factors are used for jets and the leading muon and electron in the signal and control regions. Correct scale factors are not available for the subleading muon in the produced datasets at the time of the analysis. Furthermore, two measurement regions are applied when developing the data-driven model. In these regions, two leptons are required which will be referred to as the tag and probe lepton. Scale factors are applied for the tag lepton but not for the probe lepton.

6.3 Lepton truth matching

In the analysis, the reducible background containing fake leptons is estimated using a data-driven method called the fake factor method. So MC simulations of the background processes are only used for irreducible backgrounds where the selected leptons are truly isolated leptons arising from a prompt decay (called real). This is explained in detail in Sec. 8.

If for example a $t\bar{t}W$ sample was used and 3 leptons are found passing the signal selection, then this could be satisfied by the trilepton real events with two leptons from semi-leptonic top quark decays and one from the W-boson decay, or it could be satisfied by two real leptons from top or W-boson decays and one lepton from a semi-leptonic heavy flavor decay. The former event should be included as an irreducible background event modelled by MC whereas the latter is included in the data-driven estimate and should not be included through MC. By using truth matching (i.e. assessing whether a selected lepton is originating from a true isolated lepton or a lepton in a jet) one can count correctly the irreducible components of background and ensure that events are not double counted in the data driven estimate.

MC samples are produced by combining the simulation result from an event generator such as Pythia, Sherpa or Herwig, which models the truth particles, with the tool Geant4 [59] used to simulate the passages of the particles through the detector. Objects are reconstructed based on the simulations of the detector hits. By matching the reconstructed tracks with the track hits in the Geant4 truth record, the truth origin and type of the leptons can be determined. The algorithm used to classify the truth particles is a part of the ATLAS software [60].

7 Signal Selection

The signal selection is defined in this section which is a list of requirements used to determine the signal region. First, the definition of the signal region and the additional regions used in the analysis are presented.

Three kinds of regions are applied in this analysis which are each defined by a list of kinematic requirements. The regions are called: signal region, measurement region, and control region. The signal region, which is the focus of this section, is the region which is sensitive to the signal events. It is designed such that an excess of signal events over the number of background events can be measured. No signal is expected in the measurement and control regions which are orthogonal to of the signal region, so at least one of the requirements for the signal region must be inverted when defining the requirements for the measurement and control regions. In order to ensure that the analysis is unbiased, the data in the signal region is “blinded” and can therefore not be applied until the analysis is final. The signal-free measurement and control regions are not blinded so given that they are defined similar to the signal region, results from these regions can be extrapolated to the signal region. Both the measurement and control regions are used for the background estimation model developed in Sec. 8.

The method used to choose the signal region requirements can have a big influence on the outcome. Some of the aspects which have to be considered are:

- The applied variables and their correlations.
- The statistics available and their uncertainties.
- The measure of the requirement performance, i.e. significance.

The variables applied in the signal region requirements need to have discriminating power which means that they can be used to separate signal-like and background-like events. The correlations of the applied variables also influence the selection outcome. In order to get the best signal efficiency, the variables which are best at separating signal and background events should be applied first so the interesting subset of events are considered. Furthermore, the statistics of the background processes in the region and their uncertainties should be taken into account especially in cases where the statistics are low to ensure that the signal region requirements are not optimized based on statistical fluctuations. In this analysis, the statistics in the signal region are low due to the rare topology of the signal final state. The significance should therefore be chosen such that it is stable for small numbers of background events.

The signal region requirements are determined for the five signal samples with $M_N = 5$, 10, 20, 30 and 50 GeV for which the applied mixing angles and decay lengths are listed in Tab. 3. The background processes are modeled using MC samples which are listed in Tab. 5. All the background processes are scaled to an integrated luminosity of 79.8 fb^{-1}.

In this section, the preselection requirements are first examined. These are the basic requirements to decide the subset of interesting events. Thereafter, the applied definition of the discovery significance is presented. Finally, the signal region requirements are determined.

7.1 Preselection

The preselection requirements are a list of cuts applied before the optimization of the signal region selection. In addition to the basic object selection as described in Section 6, each event is required to have exactly three leptons in the final state with the following flavor and charge composition: $\mu^{\pm}\mu^{\pm}e^{\mp}$. This focuses the search on a HNL of Majorana nature. The leading muon, which is the muon with the highest p_T, will be referred to as μ_1, and similarly the subleading muon will be referred to as μ_2. Below, the preselection requirements are examined. The final list of signal region requirements is shown in Tab. 8.

The lowest unprescaled and asymmetric di-muon trigger is used to collect the dataset. Table 9 shows the HLT name and requirements applied in the datasets from 2015-2017. In the trigger name, information is given about the trigger level, multiplicity, particle type and p_T-threshold. Further selection criteria as the tightness or type of identification and isolation will also be appended to the trigger name. For the trigger name `HLT_mu18_mu8noL1`, it is meant that it is a High Level Trigger, two muons are required in the final state with an asymmetric selection for which $p_T(\mu_1) \geq 18$ GeV and $p_T(\mu_2) \geq 8$ GeV, and the "noL1" attached to the p_T-threshold value indicates that the trigger is executed without a requirement on the L1 seed on the subleading muon. In addition to requiring the chosen trigger to be fired, it is also required that the reconstructed leading and subleading muon are the objects which caused the trigger to be fired. This is denoted trigger matching for which the reconstructed objects are matched to the trigger. In addition to the trigger matching are offline requirements made for i.e. the p_T of the matched objects. As shown in Tab. 9, one typically requires an offline p_T value which is a few GeV above the trigger threshold in order to be at full efficiency for the trigger.

Four additional requirements are included in the preselection requirements which are exploiting the final state topology of the signal:

- $p_T(\mu_2) < 50$ GeV
- $p_T(e) < 50$ GeV
- 40 GeV $< M(\mu_1, \mu_2, e) <$ 90 GeV
- $M(\mu_2, e) < M_N$

The subleading muon and electron both appear in the final state of the signal as decay products from the sterile neutrino. Therefore, $p_T(\mu_2)$ and $p_T(e)$ are naturally correlated with M_N which defines the maximum transverse momentum. Figure 7 and 8 show the distribution of $p_T(\mu_2)$ and $p_T(e)$ respectively for the signal samples with the lowest and highest mass point: $M_N = 5$ and 50 GeV. Both the distribution of $p_T(\mu_2)$ and $p_T(e)$ peak at small p_T for both the background processes and the signal but the tail of the distribution is much longer for $p_T(e)$. It is found that the requirement $p_T(\mu_2) < 50$ GeV has almost no impact on either the number of signal or background events. The requirement is still applied since an upper threshold on $p_T(\mu_2)$ is simplifying the usage of the background estimation method for fake leptons described in Sec. 8. The requirement $p_T(e) < 50$ GeV has a large impact on the number of background events which is reduced significantly. Again, the efficiencies of the signal samples are almost unaffected.

Preselection requirements	
1.	Two `loose` and `combined` muons
2.	One `LooseAndBLayerLLH` electron
3.	$\mu^{\pm}\mu^{\pm}e^{\mp}$
4.	$\lvert\eta(\mu)\rvert < 2.5$ and $\lvert\eta(e)\rvert < 2.47$ except $1.37 < \lvert\eta(e)\rvert < 1.52$
5.	$\mu : p_T \geq 5$ GeV
6.	$e : p_T \geq 7$ GeV (2015), $p_T \geq 5$ GeV (2016+2017)
7.	$\mu_1 : \lvert d_0\rvert/\sigma(d_0) < 3$ and $\lvert\Delta z_0 \sin\theta\rvert < 0.5$ mm
8.	$\mu_2 + e : \lvert\Delta z_0 \sin\theta\rvert < 1.0$ mm
9.	Trigger: `HLT_mu18_mu8noL1` (2015), `HLT_mu22_mu8noL1` (2016+2017)
10.	Trigger matching and offline p_T requirements
11.	$\mu_2 : p_T < 50$ GeV
12.	$e : p_T < 50$ GeV
13.	$40\,\text{GeV} < M(\mu_1, \mu_2, e) < 90\,\text{GeV}$
14.	$M(\mu_2, e) < 50$ GeV

Table 8: The preselection requirements applied in the signal region prior to the signal region optimization.

	Trigger	Offline requirements
2015	`HLT_mu18_mu8noL1`	$p_T(\mu_1) > 19$ GeV and $p_T(\mu_2) > 9$ GeV
2016	`HLT_mu22_mu8noL1`	$p_T(\mu_1) > 23$ GeV and $p_T(\mu_2) > 9$ GeV
2017	Same as above	Same as above

Table 9: The asymmetric di-muon trigger used to collect the dataset during the three data taking periods: 2015, 2016, and 2017.

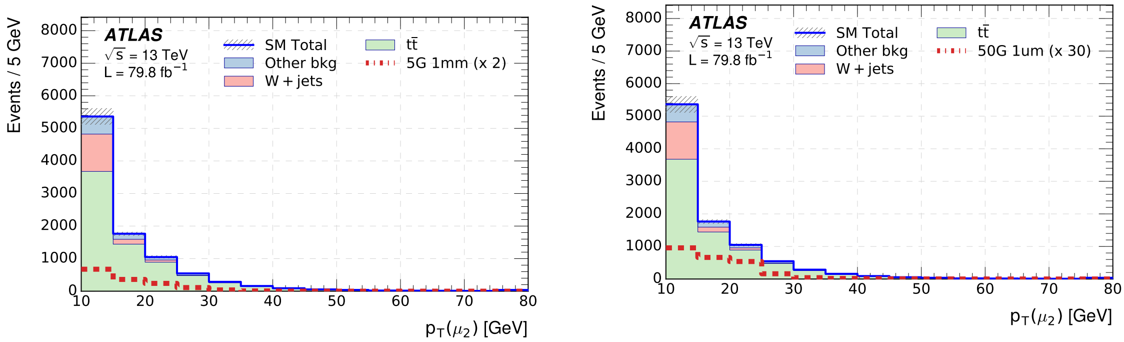
The trilepton invariant mass, $M(\mu_1, \mu_2, e)$, is a discriminating variable due to its relation to the W-boson invariant mass for the signal process. In the considered signal region channel:

$$W^{\pm} \to \mu_1^{\pm}\mu_2^{\pm}e^{\mp}\nu_e,$$

the trilepton invariant mass will yield a peak close to the mass of the W-boson at 80 GeV. Due to the neutrino in the final state of the signal process, the trilepton invariant mass will be lower than the W-boson mass. For larger M_N, the final state neutrino will in general have a larger momentum so $M(\mu_1, \mu_2, e)$ tends to be smaller when M_N is large.

The dilepton invariant mass refers to the invariant mass of the subleading muon and the electron[1] In the considered signal region channel, the invariant mass of the HNL is related to the invariant mass of the subleading muon and the electron. Due to the missing energy from the neutrino in the final state, the dilepton invariant mass will typically peak slightly below the mass of the HNL: $M(\mu_2, e) < M_N$. Here, it is assumed that the reconstructed leading muon μ_1 is the muon which first decays from the W-boson. For small M_N, this is a very good assumption since less energy is available for the leptons decaying from the HNL. For increasing M_N, the probability that the reconstructed leading muon in the signal process can be produced in the HNL decay increases.

[1] In the limit where $M(\mu_2) = M(e) = 0$ is the dilepton invariant mass:
$M(\mu_2, e) = \sqrt{[E(\mu_2) + E(e)]^2 - [\vec{p}(\mu_2) + \vec{p}(e)]^2} = \sqrt{2p_T(\mu_2)p_T(e)\,[\cosh(\eta(\mu_2) - \eta(e)) - \cos(\phi(\mu_2) - \phi(e))]}$.
It has been used that $\lvert\vec{p}\rvert = p_T \cosh\eta$, $p_x = p_T\cos\phi$, $p_x = p_T \sin\phi$ and $p_z = p_T \sinh\eta$.

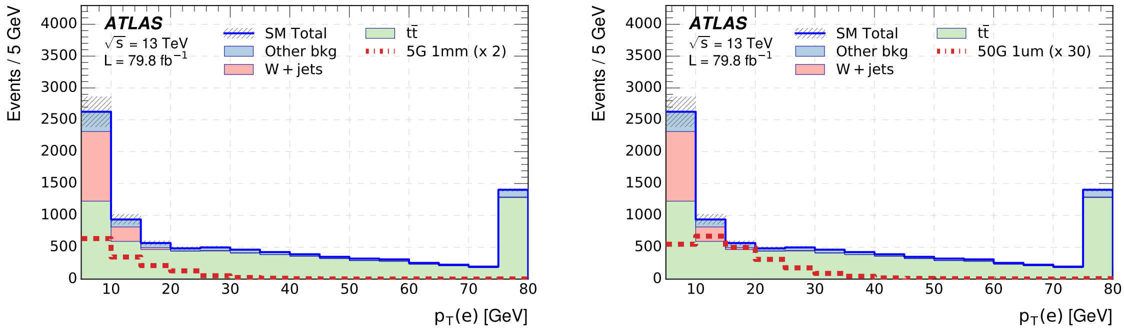


(a) $M_N = 5$ GeV.

(b) $M_N = 50$ GeV.

s

Figure 7: Signal region: $p_T(\mu_2)$ distribution after the preliminary cuts except $p_T(\mu_2) < 50$ GeV, $p_T(e) < 50$ GeV, 40 GeV $< M(\mu_1, \mu_2, e) < 90$ GeV and $M(\mu_2, e) < 50$ GeV, see Tab. 8. The blue line indicates the sum of the SM background processes and the dashed area the corresponding statistical uncertainty. The last bin in the histogram includes all events with values larger than the range of the histogram. The number of signal events in Fig. 7a and 7b are increased by a factor 2 and 30 respectively.

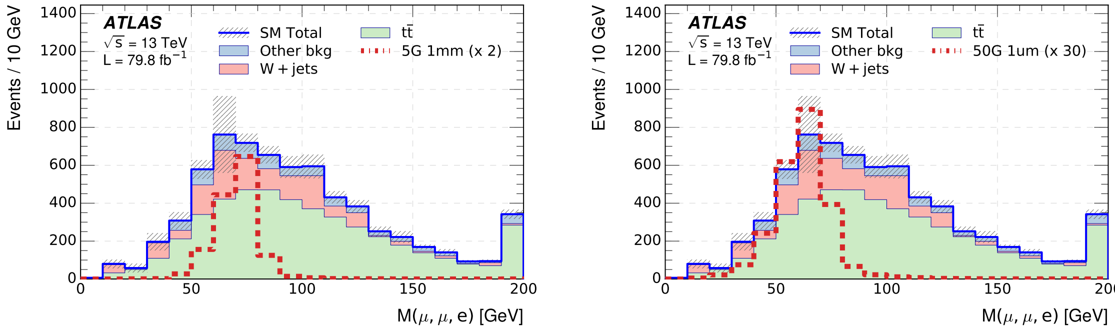


(a) $M_N = 5$ GeV

(b) $M_N = 50$ GeV

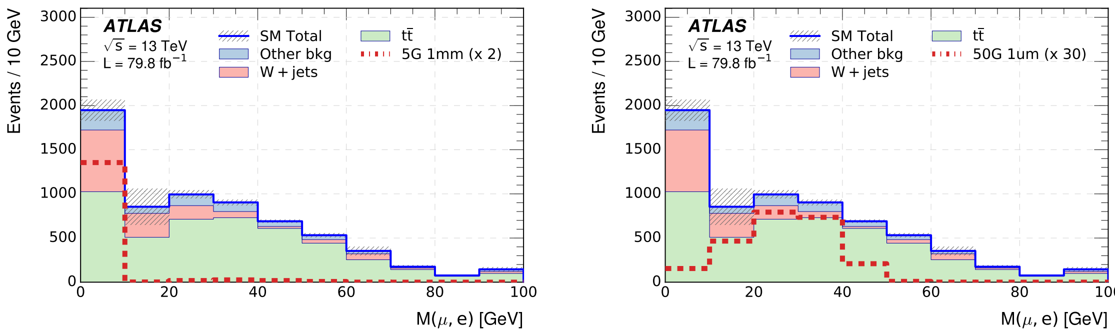
Figure 8: Signal region: $p_T(e)$ distribution after the preliminary cuts except $p_T(\mu_2) < 50$ GeV, $p_T(e) < 50$ GeV, 40 GeV $< M(\mu_1, \mu_2, e) < 90$ GeV and $M(\mu_2, e) < 50$ GeV, see Tab. 8. The blue line indicates the sum of the SM background processes and the dashed area the corresponding statistical uncertainty. The last bin in the histogram includes all events with values larger than the range of the histogram. The number of signal events in Fig. 8a and 8b are increased by a factor 2 and 30 respectively.

Figure 9 and 10 shows the distribution of $M(\mu_1, \mu_2, e)$ and $M(\mu_2, e)$ respectively for the signal samples with $M_N = 5$ and 50 GeV. When applying the requirement 40 GeV $< M(\mu_1, \mu_2, e) < 90$ GeV, the total number of background events is reduced by 55%. The requirement $M(\mu_2, e) < 50$ GeV rejects a relatively small amount of background events.



(a) $M_N = 5$ GeV. **(b)** $M_N = 50$ GeV.

Figure 9: Signal region: $M(\mu_1, \mu_2, e)$ distribution after the preliminary cuts except 40 GeV $< M(\mu_1, \mu_2, e) <$ 90 GeV and $M(\mu_2, e) <$ 50 GeV, see Tab. 8. The blue line indicates the sum of the SM background processes and the dashed area the corresponding statistical uncertainty. The last bin in the histogram includes all events with values larger than the range of the histogram. The number of signal events in Fig. 9a and 9b are increased by a factor 2 and 30 respectively.



(a) $M_N = 5$ GeV. **(b)** $M_N = 50$ GeV.

Figure 10: Signal region: $M(\mu_2, e)$ distribution after the preliminary cuts except 40 GeV $< M(\mu_1, \mu_2, e) <$ 90 GeV and $M(\mu_2, e) <$ 50 GeV, see Tab. 8. The blue line indicates the sum of the SM background processes and the dashed area the corresponding statistical uncertainty. The last bin in the histogram includes all events with values larger than the range of the histogram. The number of signal events in Fig. 10a and 10b are increased by a factor 2 and 30 respectively.

7.2 Significance

Optimizing the signal selection means maximizing the expected discovery significance for the signal. Naturally, the definition of the expected discovery significance can affect the chosen signal region requirement. The optimal definition of the significance depends on: whether the background is known, the number of background events, and how the background is determined and its related systematic uncertainty. Often, the quantities $s/\sqrt{b}$ or $s/\sqrt{b + \sigma_b^2}$ are used as measures of the expected discovery significance for a Poisson counting experiment, where s and b refer to the number of signal and background events respectively [61]. When deriving these expressions for the significance, the variance is equal to the mean for Poisson distributions, which under a background-only hypothesis gives $\mathrm{Var}[x|s = 0] = \bar{x} = b$, where x is the Poisson distributed variable. The additional uncertainty σ_b refers to the uncertainty on the background estimation given by MC statistical uncertainty and systematic uncertainty. These significance definitions rely on s/b and σ_b^2/b being small since they are determined by approximating the Poisson distribution with a Gaussian distribution and also use the assumption that the background is known.

In this analysis, the statistics is very limited. Therefore, the assumption that the number of background events is much larger than the number of signal events is not good throughout the signal optimization. The number of MC events generated in this region is also relatively small for some background samples for which it is necessary to include the statistical uncertainty related to the MC modeling: $(\sigma_b^{\text{MC stat}})^2 = \sum_{i=1}^{N_{\text{MC}}}(w_i)^2$, where w_i is the weight of the i'th MC event in the region and N_{MC} is the total number of MC events satisfying the signal region requirements. Furthermore, the background events contributing to the signal region are mainly events containing fake leptons which will be modeled using the fake factor method and thereby be subject to a not negligible systematic uncertainty, σ_b^{syst}.

The significance applied in the optimization study is defined for Poisson counting experiments with uncertain background. In the derivation it is taken into account that the expected number of background events is constrained using a control measurement. The control measurement should be made in a region where signal events are believed to be absent and where the mean number of events can be related to the number of background events in the signal region. The discovery significance determined for the background-only hypothesis given these constraints is [61]

$$Z = \left[2\left((s+b)\ln\left[\frac{(s+b)(b+\sigma_b^2)}{b^2+(s+b)\sigma_b^2}\right] - \frac{b^2}{\sigma_b^2}\ln\left[1+\frac{\sigma_b^2 s}{b(b+\sigma_b^2)}\right]\right)\right]^{1/2}, \tag{7.1}$$

for which $\sigma_b^2 = (\sigma_b^{\text{MC stat}})^2 + (\sigma_b^{\text{syst}})^2$. When expanding the significance in powers of s/b and σ_b^2/b it is found that [61] $Z = s/\sqrt{b+\sigma_b^2}(1+\mathcal{O}(s/b)+\mathcal{O}(\sigma_b^2/b)$, which is exactly the discovery significances described above. The significance as determined using Eq. 7.1 is evaluated using the total amount of background and signal events in the region as a one-bin significance. In this analysis, the data in the signal region will be compared to the estimated background and signal contribution using a binned distribution. By binning the significance, the uncertainties of the background samples are properly taken into account. Therefore, in the optimization study of this analysis, the significance is determined as:

$$Z^2 = \sum_{j=1}^{N_{\text{bins}}}\left[2\left((s_j+b_j)\ln\left[\frac{(s_j+b_j)(b_j+\sigma_{b,j}^2)}{b_j^2+(s_j+b_j)\sigma_{b,j}^2}\right] - \frac{b_j^2}{\sigma_{b,j}^2}\ln\left[1+\frac{\sigma_{b,j}^2 s_j}{b_j(b_j+\sigma_{b,j}^2)}\right]\right)\right], \tag{7.2}$$

for which s_j, b_j and $(\sigma_{b,j})^2 = (\sigma_{b,j}^{\text{MC stat}})^2 + (\sigma_{b,j}^{\text{syst}})^2$ are determined using a sum over the event weights which populate the j'th bin. For a systematic uncertainty on the background estimation defined in terms of the number of background events, the total systematic uncertainty of the background is:

$$\sigma_b^{\text{syst}} = b \cdot p_{\text{syst}}. \tag{7.3}$$

Systematic uncertainties can also be assigned for a subset of the background events for which p_{syst} will not be a constant but depend on the relevant variables.

7.3 Signal region optimization study

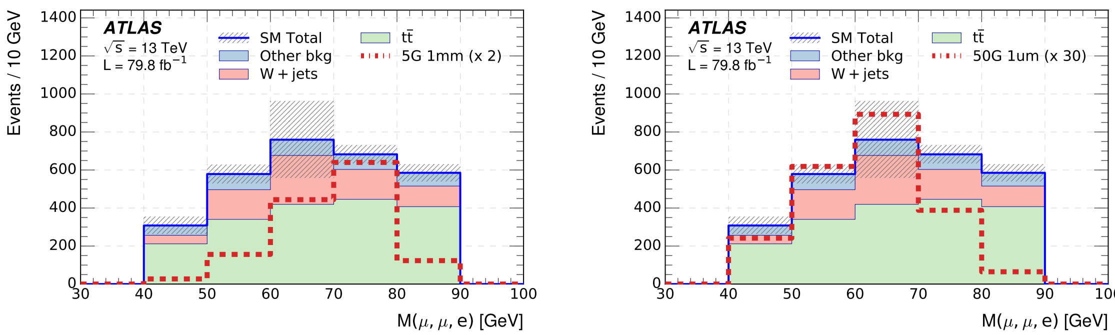
In this section, the final list of signal region requirements are determined. Prior to the study are the preselection requirements listed in Tab. 8 applied. The following variables are considered in the signal selection:

1. $p_T(\mu_1)$
2. $p_T(e)$
3. $p_T(\mu_2)$
4. $N_{\text{b-jets}}$
5. Isolation for μ_1
6. Isolation for e
7. Isolation for μ_2
8. E_T^{miss}
9. Identification for μ_1
10. $\Delta R(\mu_2, e)$
11. $|d_0|/\sigma(d_0)(e)$
12. $|d_0|/\sigma(d_0)(\mu_2)$
13. $|\Delta z_0 \sin(\theta)|(e)$
14. $|\Delta z_0 \sin(\theta)|(\mu_2)$

After a signal region requirement has been chosen for a variable, it will be included in the signal region requirements and applied prior to the optimization study of the next variable.

The optimal requirements are based on the significance as described in the previous section, see

Eq. 7.2, where the $M(\mu_1, \mu_2, e)$ distribution is used to bin the significance. Figure 11 shows the $M(\mu_1, \mu_2, e)$ distribution for the two signal samples with $M_N = 5$ and 50 GeV. A study of the binned significance is shown in appendix Sec. B. In the study, it is found that the variable $M(\mu_1, \mu_2, e)$ is a good variable to apply because it discriminates signal-like events from background-like events due to the correlation between the W-boson mass and M_N for the signal samples. The bins used as a function of the $M(\mu_1, \mu_2, e)$ variable are defined in the range from 40-90 GeV for which the size of each bin is: $(50/N_{\text{bins}})$ GeV. The significance is determined for both $N_{\text{bins}} = 5$ and 10. In regions with high statistics, the number of bins will have a minimal effect on the estimated significance. When fewer background events are present, the number of bins have a large impact on the relative size of the background uncertainty in each bin. Since only statistical uncertainties are used for the MC samples, the following systematic uncertainties are applied: $p_{\text{syst}} = 0.3$ and 0.5 corresponding to 30% and 50% uncertainty. In section 8, a background estimation method is developed using the fake factor method for which other analyses applying this method have estimated systematic uncertainties in this range. By examining the estimated significance for all the signal samples as well as the different combinations of N_{bins} and p_{syst}, it is both checked that the optimal cut performs well for all the signal samples and that it is robust against statistical differences.



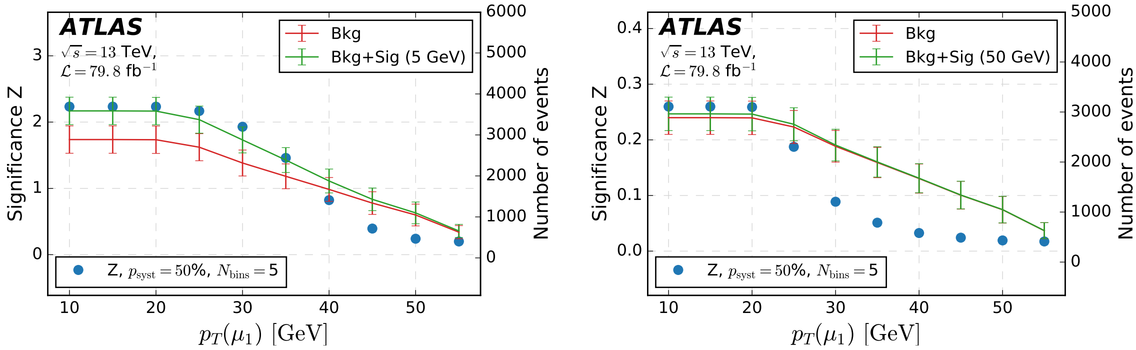
(a) $M_N = 5$ GeV.

(b) $M_N = 50$ GeV.

Figure 11: Signal region: $M(\mu_1, \mu_2, e)$ distribution after applying the preselection requirements, see Tab. 8. The blue line indicates the sum of the SM background processes and the dashed area the corresponding statistical uncertainty. The last bin in the histogram includes all events with values larger than the range of the histogram. The number of signal events in Fig. 10a and 10b are increased by a factor 2 and 30 respectively.

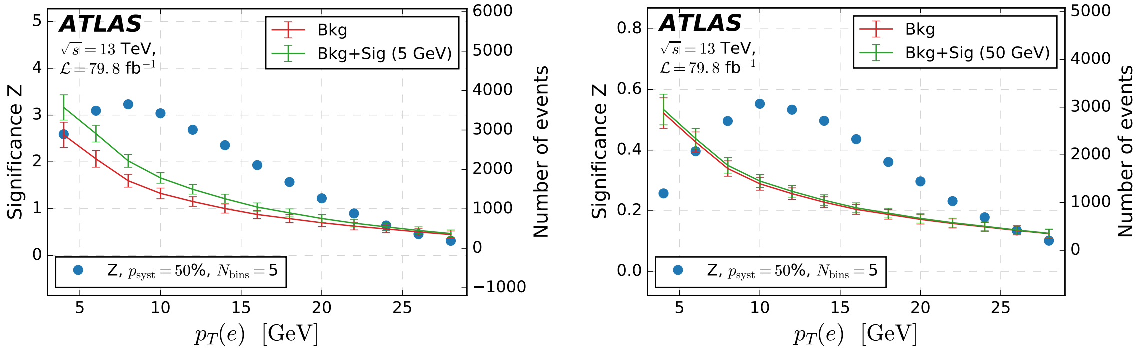
In this section, I have only included a small number of the plots considered in the optimization study in order to motivate the chosen signal region requirements without having to include the full study. Plots showing the significance determined using $N_{\text{bins}} = 5$ and $p_{\text{syst}} = 0.5$ are included for the two extreme mass points $M_N = 5$ and 50 GeV. In appendix Sec. C, the optimization study plots are shown for $M_N = 5$, 20 and 50 GeV with all the four combinations of $N_{\text{bins}} = 5/10$ and $p_{\text{syst}} = 0.3/0.5$. The plots for the signal samples with mass $M_N = 10$ and 30 GeV are not included since the preferred cuts for these samples lies in between the preferred values for their respective "neighboring" mass points. When comparing the significance determined for the four combinations of $N_{\text{bins}} = 5/10$ and $p_{\text{syst}} = 0.3/0.5$ it is found that in the optimization study of the first variables, where the statistics of both the signal and background samples are relatively large, the shape of the significance is unchanged. The largest significance is determined for $N_{\text{bins}} = 10$ with the lowest systematic uncertainty of $p_{\text{syst}} = 0.3$. When the number of background events is reduced to a couple of hundred events, variations in the significance shape are found especially when changing the number of bins. The requirements which are chosen in the optimization study perform well across the four combinations.

The first variable which is considered is $p_T(\mu_1)$. Figure 12 shows the estimated significance and number of background and signal events achieved when applying different cuts on the variable $p_T(\mu_1)$ for the signal samples with $M_N = 5$ and 50 GeV. Given a value x_0 on the x-axis, the



(a) M_N = 5 GeV **(b)** M_N = 50 GeV

Figure 12: Optimization study for the variable $p_T(\mu_1)$ for M_N = 5 and 50 GeV with $p_{\rm syst}$ = 50% and $N_{\rm bins}$ = 5. The plots show the significance and number of events as a function of the cut on $p_T(\mu_1)$. The x-axis value x_0 represents the requirement $p_T(\mu_1) \geq x_0$ GeV. The green and red line represent the number of background events and the sum of the number of background and signal events respectively.



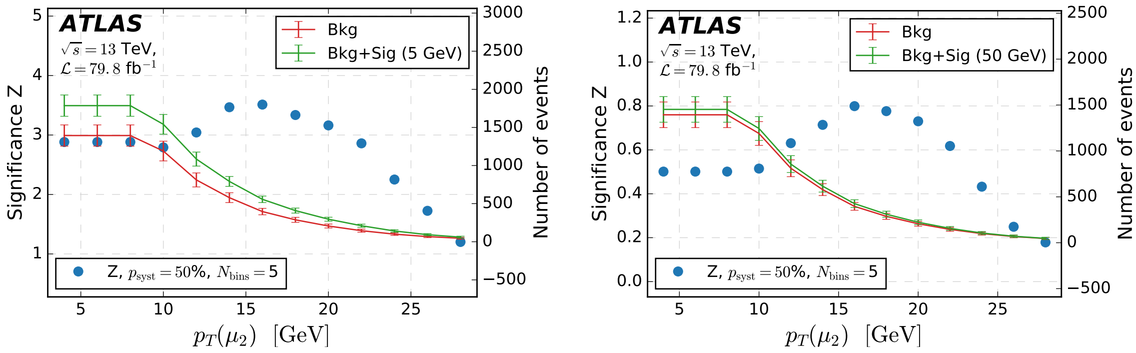
(a) M_N = 5 GeV **(b)** M_N = 50 GeV

Figure 13: Optimization study for the variable $p_T(e)$ for M_N = 5 and 50 GeV with $p_{\rm syst}$ = 50% and $N_{\rm bins}$ = 5. The plots show the significance and number of events as a function of the cut on $p_T(e)$. The x-axis value x_0 represents the requirement $p_T(e) \geq x_0$ GeV. The green and red line represent the number of background events and the sum of the number of background and signal events respectively.

corresponding significance and number of events are determined using the extra signal region requirement: $p_T(\mu_1) \geq x_0$ GeV. It is found that the best cut for all the signal samples is $p_T(\mu_1) \geq 23$ GeV, which is the lowest possible offline $p_T(\mu_1)$ requirement for the applied High Level Trigger labeled `HLT_mu22_mu8noL1` (active for 2016 and 2017 data).

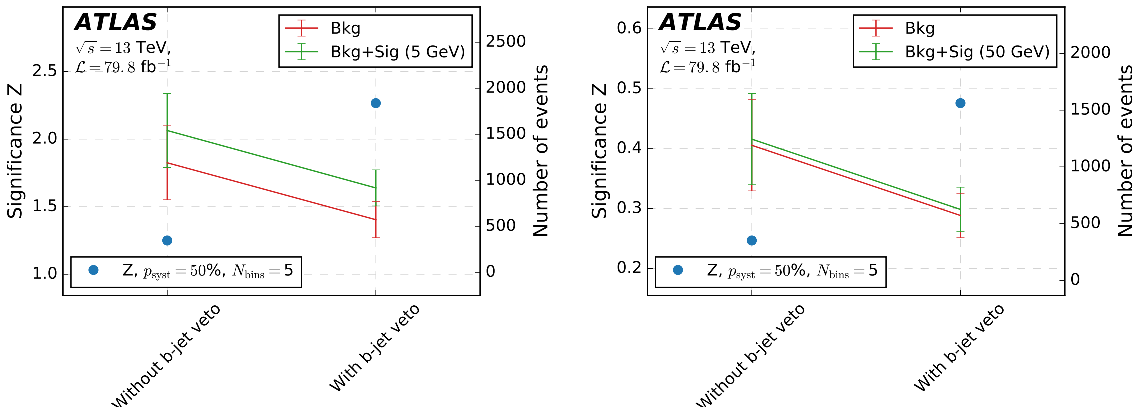
The second variable considered is $p_{\rm T}(e)$. Figure 13 shows the significance scan of the variable in the same way as for the variable $p_T(\mu_1)$ above. Again, plots are included for the signal samples with M_N = 5 and 50 GeV for $N_{\rm bins}$ = 5 and $p_{\rm syst}$ = 0.5. For the signal samples with small M_N, $p_{\rm T}(e)$ is generally lower than for the samples with higher M_N since the electron is created after the HNL decay: $N \rightarrow \mu_2^{\pm} e^{\mp} \nu_e$. For all the signal samples, the significance is peaking around $p_{\rm T}(e) \geq 10$ GeV which is chosen as a signal region requirement.

Due to the asymmetric dimuon trigger applied in the signal region, an offline requirement is applied to the subleading muon of $p_T(\mu_2) \geq 9$ GeV. Figure 14 shows the significance scan for the signal samples with $M_N = 5$ and 50 GeV. As for the optimization study of the variable $p_{\rm T}(e)$, it is shown that the value of $p_T(\mu_2)$ is correlated with the mass point M_N. The significance is peaking when applying the requirement $p_T(\mu_2) \geq 14 - 16$ GeV even though this requirement would decrease the efficiency for the signal samples with small M_N dramatically. A looser signal requirement is chosen despite the possible gain in significance since variables as the isolation of μ_2 is correlated



(a) $M_N = 5$ GeV **(b)** $M_N = 50$ GeV

Figure 14: Optimization study for the variable $p_T(\mu_2)$ for $M_N = 5$ and 50 GeV with $p_{\rm syst} = 50\%$ and $N_{\rm bins} = 5$. The plots show the significance and number of events as a function of the cut on $p_T(\mu_2)$. The x-axis value x_0 represents the requirement $p_T(\mu_2) \geq x_0$ GeV. The green and red line represent the number of background events and the sum of the number of background and signal events respectively.



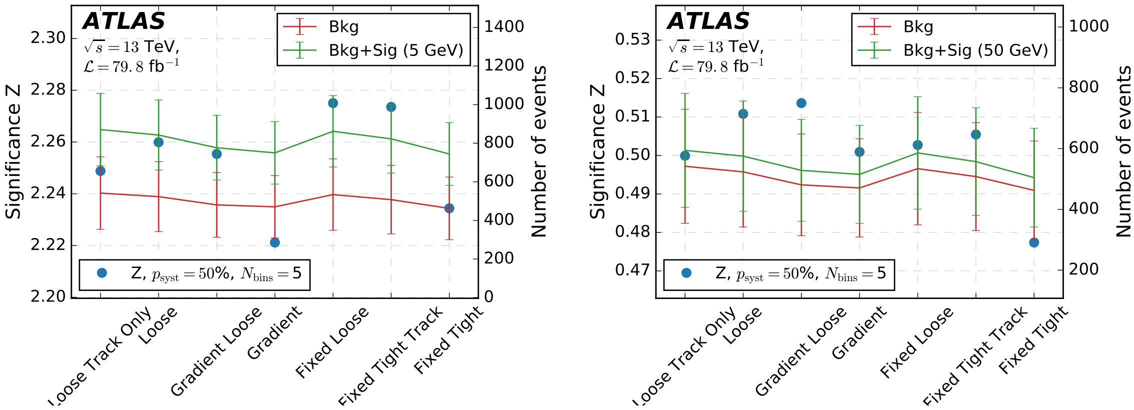
(a) $M_N = 5$ GeV **(b)** $M_N = 50$ GeV

Figure 15: Optimization study for the variable $N_{\text{b-jets}}$ for $M_N = 5$ and 50 GeV with $p_{\rm syst} = 50\%$ and $N_{\rm bins} = 5$. The plots show the significance and number of events as a function of the cut on $N_{\text{b-jets}}$. The green and red line represent the number of background events and the sum of the number of background and signal events respectively.

with $p_T(\mu_2)$ and is expected to perform better at discriminating signal-like and background-like events. The signal requirement $p_T(\mu_2) \geq 10$ GeV is chosen.

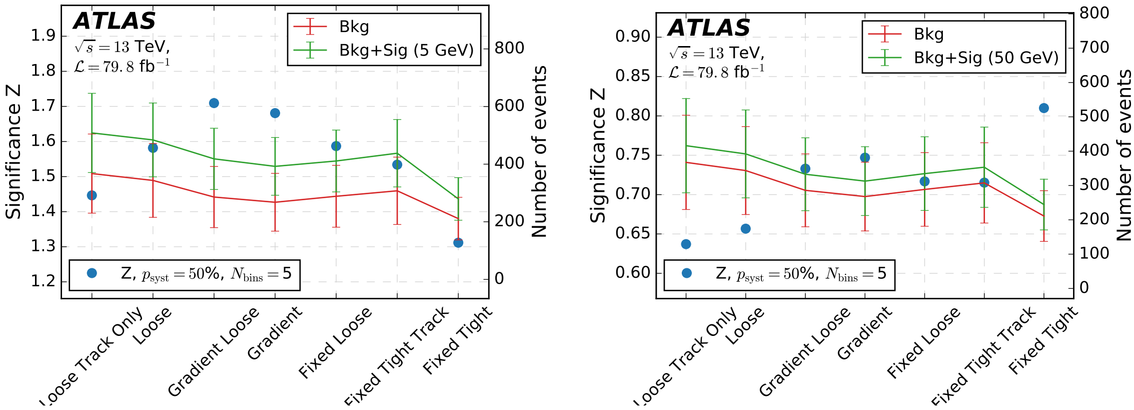
The effect of applying a b-jet veto is then examined. Figure 15 shows the significance and number of events in the signal region before and after applying a b-jet veto. Since there is no b-jets present in the final state of the signal process, the b-jet veto is only decreasing the number of background events which increases the significance. So the requirement $N_{\text{b-jets}} = 0$ is used in the signal region.

The optimization of the isolation selections applied for the three final state leptons are then examined. In Sec. 6.1.6, the available isolation selections were presented. Figure 16 shows the optimization study for the different isolation selections for μ_1. For the signal samples, the leading muon is in general well-isolated since it arise from the first W-boson decay. Also for the background samples, the leading muon is in general a prompt, well-isolated muon. Therefore, the significance is relatively constant for all the isolation selections. The fixed cut isolation denoted `FixedCutTightTrackOnly` is used for μ_1 since it is a hard enough cut to remove some of the background events for which the leading muon is a fake or non-prompt muon while rejecting very few signal events.



(a) $M_N = 5$ GeV **(b)** $M_N = 50$ GeV

Figure 16: Optimization study for the isolation selection of μ_1 for $M_N = 5$ and 50 GeV with $p_{\rm syst} = 50\%$ and $N_{\rm bins} = 5$. The plots show the significance and number of events as a function of the isolation of μ_1. The green and red line represent the number of background events and the sum of the number of background and signal events respectively.



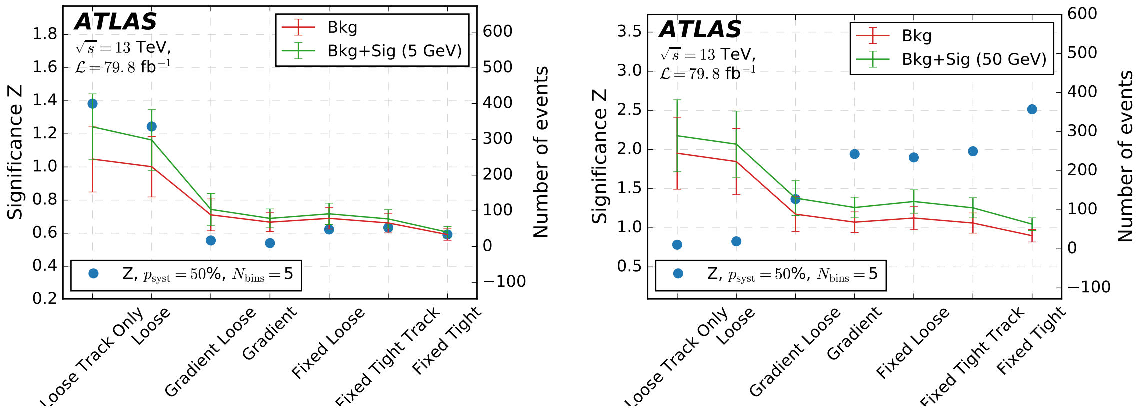
(a) $M_N = 5$ GeV **(b)** $M_N = 50$ GeV

Figure 17: Optimization study for the isolation selection of e for $M_N = 5$ and 50 GeV with $p_{\rm syst} = 50\%$ and $N_{\rm bins} = 5$. The plots show the significance and number of events as a function of the isolation of e. The green and red line represent the number of background events and the sum of the number of background and signal events respectively.

The optimization study for the isolation selection of e is shown in Fig. 17. For signal samples with small M_N, looser isolation selections are preferred compared to the signal samples with higher M_N. When the HNL mass is small, the HNL is in general more boosted and therefore its decay products tend to be more collimated and thereby less isolated. The efficiency based isolation named `Gradient` is chosen for e since it performs well for all the signal samples.

The optimization study for the isolation selection of μ_2 is shown in Fig. 18. It is shown that the two loosest isolation selections `LooseTrackOnly` and `Loose` are very bad at discriminating fake and real leptons since very few background events are being rejected by those isolation selections even though many of the background events contain a fake subleading muon. Similarly to the study of the isolation for e above, the looser isolation selections are preferred when applying signal samples with low M_N while the tighter isolation selections are preferred when the HNL mass is larger. The isolation `FixedCutLoose` is used for μ_2 since it performs well for the majority of the signal samples and a relatively large number of background events are being kept. This choice of isolation is reducing the significance achieved for the signal sample with $M_N = 5$ GeV significantly. It is still chosen to apply the isolation `FixedCutLoose` due to the following two reasons. The prompt HNL analysis is not ideal to probe the signal from the more displaced signal samples with $M_N = 5$

and 10 GeV which have a proper decay length of $c\tau = 1$ mm, so it is not beneficial to optimize the signal region requirement for these signal samples. Furthermore, in order to successfully apply the fake factor method to estimate the fake background contribution, it is necessary to apply an isolation selection which is more effective at discriminating real and fake muons. In Sec. 8, the effect of applying the isolation selection `Loose` for the subleading muon is also studied to test the stability of the fake background model using increased statistics.



(a) $M_N = 5$ GeV **(b)** $M_N = 50$ GeV

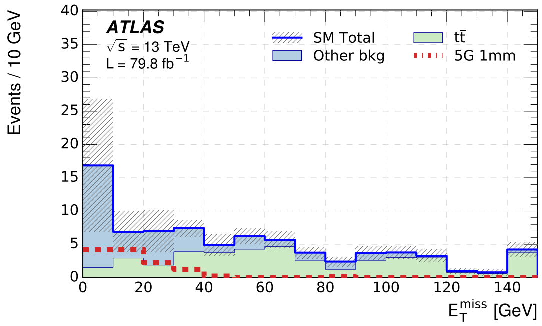
Figure 18: Optimization study for the isolation selection of μ_2 for $M_N = 5$ and 50 GeV with $p_{\text{syst}} = 50\%$ and $N_{\text{bins}} = 5$. The plots show the significance and number of events as a function of the isolation of μ_2. The green and red line represent the number of background events and the sum of the number of background and signal events respectively.

The next variable which is considered is the missing transverse energy, E_T^{miss}. Figure 19 shows the E_T^{miss} distribution in the signal region for $M_N = 5$ and 50 GeV. It is found that for all the signal samples $E_T^{\text{miss}} < 60$ GeV. Fig. 20 shows the optimization study for the E_T^{miss} variable where the point x_0 on the x-axis represents the signal region requirement $E_T^{\text{miss}} \leq x_0$ GeV. It is found that $E_T^{\text{miss}} < 60$ GeV is a good cut since no signal events are lost while the number of background events from especially $t\bar{t}$ is decreased.

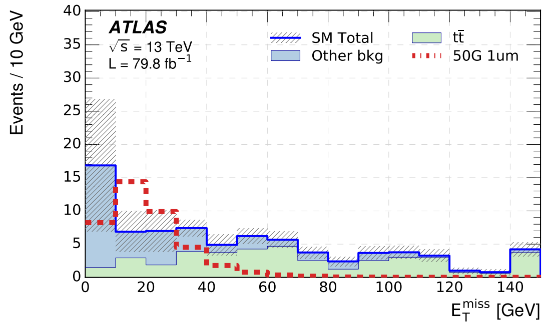
An optimization study is also made for the identification quality selection of the leading muon. It shows that there is a minimal effect when changing the identification selection of the leading muon from the loose selection to the tight selection. The plots of the optimization study are included in appendix Fig. 75-77. This shows that the leading muon is in general a good quality real muon. The tight identification selection is applied for the leading muon in order to decrease the number of events where the leading muon is not real for the background estimation model, see Sec. 8.

The variable $\Delta R(\mu_2, e) = \sqrt{(\phi(\mu_2) - \phi(e))^2 + (\eta(\mu_2) - \eta(e))^2}$ is now examined. The motivation behind making a requirement on this variable is coming from the background estimation, see Sec. G. It is found that multi-fake background, where both the subleading muon and electron are fake leptons, is not well modeled in the cases where $\Delta R(\mu_2, e)$ is small. Therefore, it is necessary to define a minimum ΔR separation which will not decrease the signal efficiency too much. Figure 21 shows the significance and number of events for the signal samples with $M_N = 5$ and 50 GeV given a range of different requirements on $\Delta R(\mu_2, e)$. It is found that the angular separation between the subleading muon and electron is in general smallest for the signal samples with low mass point M_N due to the HNL being more boosted and hence its decay products being more collimated. For the signal samples with $M_N = 10$, 20, 30 and 50 GeV, no events have $\Delta R(\mu_2, e) < 0.2$. The signal sample with $M_N = 5$ GeV has 0.5 events with $\Delta R(\mu_2, e) < 0.2$, so the number of events decrease from 6.5±0.9 events to 6.0±0.9 events. A signal region requirement is made of $\Delta R(\mu_2, e) > 0.2$.

The last variables studied are the impact parameter variables, see Sec. 6.1.7. For the subleading muon and electron, the recommended IP requirements are not applied in the preselection in order

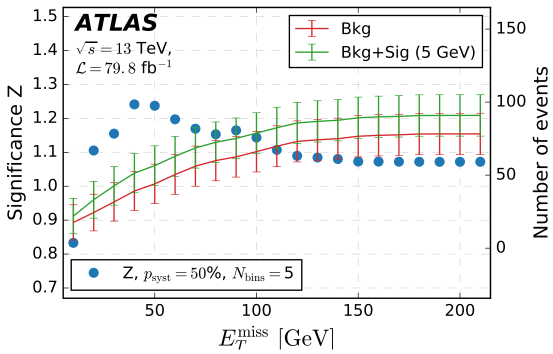


(a) $M_N = 5$ GeV

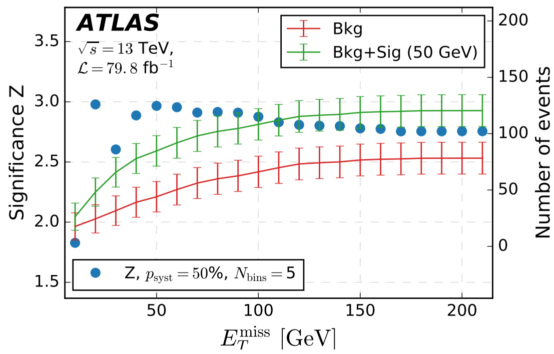


(b) $M_N = 50$ GeV

Figure 19: Distribution of E_T^{miss} in the signal region. The following signal region requirements are applied: preselection, see Tab. 8; $p_T(\mu_1) \geq 23$ GeV; $p_T(e) \geq 10$ GeV; $p_T(\mu_2) \geq 10$ GeV; $N_{\text{b-jets}} = 0$; and isolation selections `FixedCutTightTrackOnly`, `Gradient` and `FixedCutLoose` for μ_1, e and μ_2 respectively. The blue line indicates the sum of the SM background processes and the dashed area the corresponding statistical uncertainty. The last bin in the histogram includes all events with values larger than the range of the histogram.

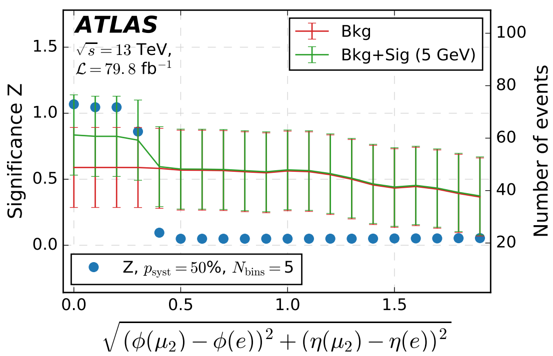


(a) $M_N = 5$ GeV

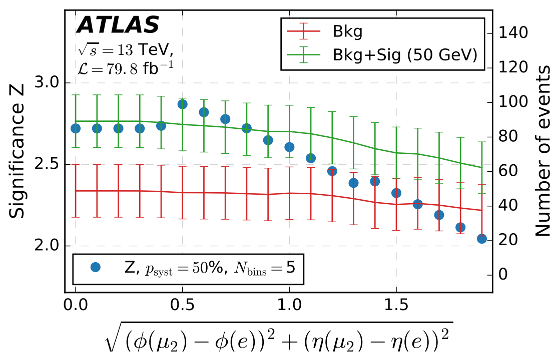


(b) $M_N = 50$ GeV

Figure 20: Optimization study for the variable E_T^{miss} for $M_N = 5$ and 50 GeV with $p_{\text{syst}} = 50\%$ and $N_{\text{bins}} = 5$. The plots show the significance and number of events as a function of the cut on E_T^{miss}. The x-axis value x_0 represents the requirement $E_T^{\text{miss}} \leq x_0$ GeV. The green and red line represent the number of background events and the sum of the number of background and signal events respectively. The last bin in the histogram includes all events with values larger than the range of the histogram.



(a) $M_N = 5$ GeV



(b) $M_N = 50$ GeV

Figure 21: Optimization study for variable $\Delta R(\mu_2, e)$ for $M_N = 5$ and 50 GeV with $p_{\text{syst}} = 50\%$ and $N_{\text{bins}} = 5$. The plots show the significance and number of events as a function of the cut on $\Delta R(\mu_2, e)$. The x-axis value x_0 represents the requirement $\Delta R(\mu_2, e) \geq x_0$ GeV. The green and red line represent the number of background events and the sum of the number of background and signal events respectively.

Variable	Signal region requirement	Figures of optimization study
$p_T(\mu_1)$	$p_T(\mu_1) \geq 23$ GeV	Fig. 12, appendix Fig. 51-53
$p_T(e)$	$p_T(e) \geq 10$ GeV	Fig. 13, appendix Fig. 54-56
$p_T(\mu_2)$	$p_T(\mu_2) \geq 10$ GeV	Fig. 14, appendix Fig. 57-59
$N_{\text{b-jets}}$	b-jet veto: $N_{\text{b-jets}} = 0$	Fig. 15, appendix Fig. 60-62
Isolation for μ_1	`FixedCutTightTrackOnly`	Fig. 16, appendix Fig. 63-65
Isolation for e	`Gradient`	Fig. 17, appendix Fig. 66-68
Isolation for μ_2	`FixedCutLoose`	Fig. 18, appendix Fig. 69-71
E_T^{miss}	$E_T^{\text{miss}} < 60$ GeV	Fig. 20, appendix Fig. 72-74
Identification for μ_1	`Tight`	Appendix Fig. 75-77
$\Delta R(\mu_2, e)$	$\Delta R(\mu_2, e) > 0.2$	Fig. 21, appendix Fig. 78-80
$\lvert d_0\rvert/\sigma(d_0)(e)$	$\lvert d_0\rvert/\sigma(d_0)(e) < 10$	Appendix Fig. 81-83
$\lvert d_0\rvert/\sigma(d_0)(\mu_2)$	$\lvert d_0\rvert/\sigma(d_0)(\mu_2) < 10$	Appendix Fig. 84-86
$\lvert\Delta z_0 \sin(\theta)\rvert(e)$	$\lvert\Delta z_0 \sin(\theta)\rvert(e) < 0.5$ mm	Appendix Fig. 87-89
$\lvert\Delta z_0 \sin(\theta)\rvert(\mu_2)$	$\lvert\Delta z_0 \sin(\theta)\rvert(\mu_2) < 0.5$ mm	Appendix Fig. 90-92

Table 10: The list of variables examined in the optimization study of the signal region requirements. Prior to the optimization study, the preselection requirements listed in Tab. 8 are applied. The signal region requirements are listed in chronological order, in the sense that prior to the optimization study of a variable both the preselection requirements and all the signal region requirements listed before the study are applied.

to avoid a possible decrease in the signal efficiency especially for the most displaced mass point with a proper decay length of $c\tau = 1$ mm.
The variable $|d_0|/\sigma(d_0)$ is first considered. The optimization scan of the variables are shown in the appendix: Fig. 81-83 for $|d_0|/\sigma(d_0)(e)$ and Fig. 84-86 for $|d_0|/\sigma(d_0)(\mu_2)$. It is found, that following the recommendations of $|d_0|/\sigma(d_0)(e) < 3$ and $|d_0|/\sigma(d_0)(\mu) < 5$ for the electron and subleading muon will decrease the number of signal events. Therefore, a looser upper threshold is applied as a signal region requirement: $|d_0|/\sigma(d_0)(e) < 10$ and $|d_0|/\sigma(d_0)(\mu_2) < 10$ for which the effect on the signal efficiency is minimal. An upper threshold is still defined so fake muons and electrons with very large $|d_0|/\sigma(d_0)$ cannot contribute to the signal region.
The variable $|\Delta z_0 \sin(\theta)|$ is now considered. The optimization scan of the variables are shown in the appendix: Fig. 87-89 for $|\Delta z_0 \sin(\theta)|(e)$ and Fig. 90-92 for $|\Delta z_0 \sin(\theta)|(\mu_2)$. It is found that the requirements $|\Delta z_0 \sin(\theta)|(e) < 0.5$ mm and $|\Delta z_0 \sin(\theta)|(\mu_2) < 0.5$ mm have a minimal impact on the number of signal events for all the five signal samples. These are the recommended IP requirements for $|\Delta z_0 \sin(\theta)|$ for both muons and electrons which is primarily used to reject fake leptons from pile-up and muons from cosmic-ray interactions.

Table 10 shows a list of all the variables which were examined in the optimization study in this section together with the signal region requirements determined doing the study and the figure numbers for the optimization study of each variable. All the signal region requirements, including the preselection, are listed in Tab. 11.

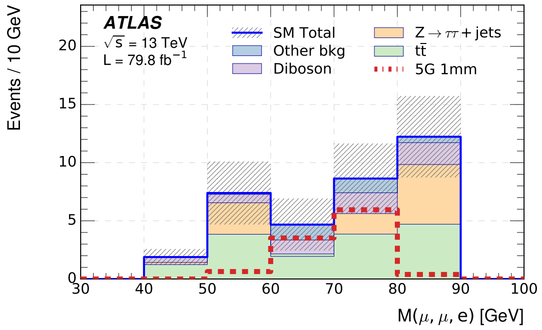
The expected sensitivity of the HNL signal is shown in Tab. 12 for the five signal samples with $M_N = 5$, 10, 20, 30 and 50 GeV. It is shown that the lowest significance is achieved for the two lowest mass points, $M_N = 5$ and 10 GeV, which are the most displaced. The highest significance is determined for the signal sample with $M_N = 30$ GeV and the proper decay length $c\tau = 10\mu m$. The expected mixing angle sensitivities are also shown which are determined for a 2σ significance of the background-only hypothesis. When estimating the mixing angle sensitivities, it is assumed that the shape of the signal sample in the five applied bins will not change. Changing the mixing angle will also change the proper decay length of the HNL signal which is not taken into account either. The mixing angle is increasing for the most displaced signal samples which corresponds to a lower proper decay length and therefore an expected increase for the signal efficiency in this prompt search.

Preselection requirements	
1.	Two `loose` and `combined` muons
2.	One `LooseAndBLayerLLH` electron
3.	$\mu^{\pm}\mu^{\pm}e^{\mp}$
4.	$\lvert\eta(\mu)\rvert < 2.5$ and $\lvert\eta(e)\rvert < 2.47$ except $1.37 < \lvert\eta(e)\rvert < 1.52$
5.	$\mu : p_T \geq 4$ GeV
6.	$e : p_T \geq 7$ GeV (2015), $p_T \geq 4.5$ GeV (2016+2017)
7.	$\mu_1 : \lvert d_0\rvert/\sigma(d_0) < 3$ and $\lvert\Delta z_0 \sin\theta\rvert < 0.5$ mm
8.	$\mu_2 + e : \lvert\Delta z_0 \sin\theta\rvert < 1.0$ mm
9.	Trigger: `HLT_mu18_mu8noL1` (2015), `HLT_mu22_mu8noL1` (2016+2017)
10.	Trigger matching and offline p_T requirements
11.	$\mu_2 : p_T < 50$ GeV
12.	$e : p_T < 50$ GeV
13.	$40\,\text{GeV} < M(\mu_1, \mu_2, e) < 90\,\text{GeV}$
14.	$M(\mu_2, e) < 50$ GeV
Optimization study requirements	
15.	$p_T(\mu_1) > 23$ GeV
16.	$p_T(\mu_2) > 10$ GeV
17.	$p_T(e) > 10$ GeV
18.	$N_{\text{b-jets}} = 0$
19.	Isolation for μ_1: `FixedCutTightTrackOnly`
20.	Isolation for μ_2: `FixedCutLoose`
21.	Isolation for e: `Gradient`
22.	$E_T^{\text{miss}} < 60$ GeV
23.	Identification selection for μ_1: `Tight`
24.	$\Delta R(\mu_2, e) > 0.2$
25.	$\lvert d_0\rvert/\sigma(d_0)(\mu_2) < 10$ and $\lvert\Delta z_0 \sin(\theta)\rvert(\mu_2) < 0.5$ mm
26.	$\lvert d_0\rvert/\sigma(d_0)(e) < 10$ and $\lvert\Delta z_0 \sin(\theta)\rvert(e) < 0.5$ mm

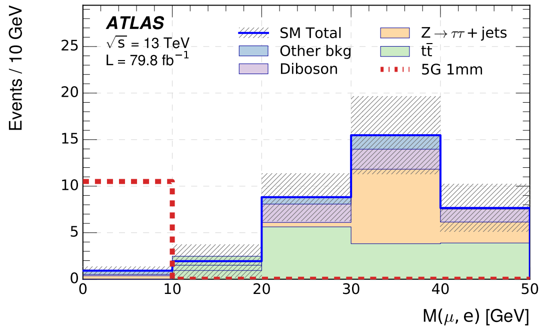
Table 11: The final list of signal region requirements.

The number of events at each stage of the signal region requirement list in Tab. 11 is shown in appendix Sec. D. Table 66 shows the number of events in the signal samples with $M_N = 5$, 10, 20, 30 and 50 GeV at each stage of the signal region requirements. Similarly, Tab. 67 and 68 show the number of events for the different background processes and the total number of background events at each stage of the signal region requirements. The contributions from the following background processes are shown: $t\bar{t}$, W+jets, Z+jets, t and Wt, triboson $t\bar{t}V$, $V\gamma^*$ and diboson events.

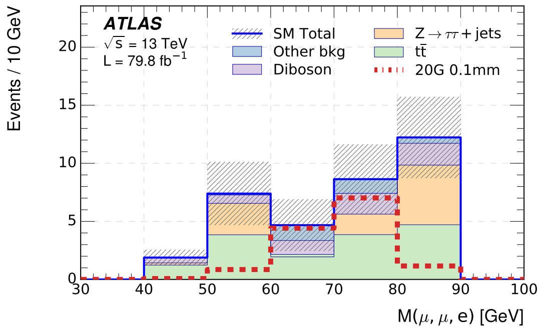
The distributions of $M(\mu_1, \mu_2, e)$ and $M(\mu_2, e)$ are shown in Fig. 22 for the signal samples with $M_N = 5$, 20 and 50 GeV after applying all the signal region requirements as listed in Tab. 11. It is shown, that the majority of the background events arise from the $t\bar{t}$ process. The second and third largest contributions are $Z \to \tau^{\pm}\tau^{\mp}$ and diboson events. The $t\bar{t}$ and Drell-Yan processes can only contribute to the signal region if one or more non-prompt or fake leptons are present in the final state. In the following section, the background estimation will be examined where focus will be on how to estimate the number of background events which contain non-prompt of fake leptons.



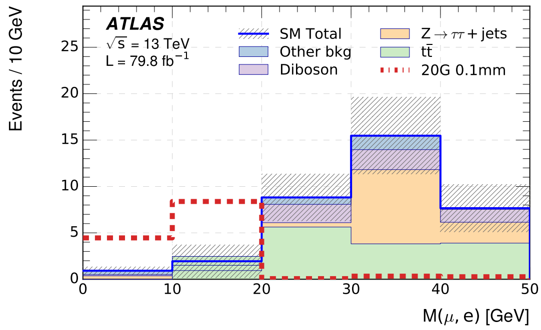
(a) $M_N = 5$ GeV. Distribution $M(\mu_1, \mu_2, e)$.



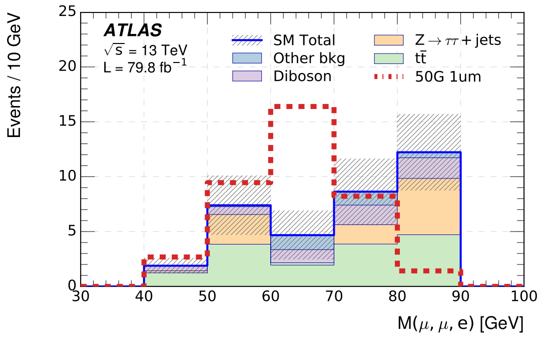
(b) $M_N = 5$ GeV. Distribution $M(\mu_2, e)$.



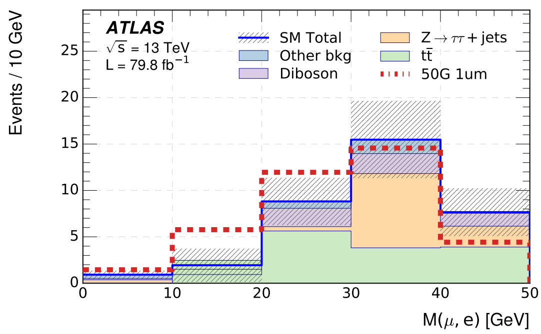
(c) $M_N = 20$ GeV. Distribution $M(\mu_1, \mu_2, e)$.



(d) $M_N = 20$ GeV. Distribution $M(\mu_2, e)$.



(e) $M_N = 50$ GeV. Distribution $M(\mu_1, \mu_2, e)$.



(f) $M_N = 50$ GeV. Distribution $M(\mu_2, e)$.

Figure 22: Distribution of $M(\mu_1, \mu_2, e)$ and $M(\mu_2, e)$ in the signal region for the signal samples with $M_N = 5$, 20 and 50 GeV. All the defined signal region requirements are applied as listed in Tab. 11. The blue line indicates the sum of the SM background processes and the dashed area the corresponding statistical uncertainty. The last bin in the histogram includes all events with values larger than the range of the histogram.

	Expected sensitivity of the HNL signal Background modeled using MC				
	$M = 5$ GeV	$M = 10$ GeV	$M = 20$ GeV	$M = 30$ GeV	$M = 50$ GeV
Mixing angle:	$3.9 \cdot 10^{-4}$	$1.2 \cdot 10^{-5}$	$3.8 \cdot 10^{-6}$	$4.6 \cdot 10^{-6}$	$3.4 \cdot 10^{-6}$
Significance Z with $N_{\text{bins}} = 5$, $\text{p}^{\text{syst}} = 50\%$	1.18	0.68	1.41	3.72	3.38
Mixing angle:	$7.4 \cdot 10^{-4}$	$4.2 \cdot 10^{-5}$	$5.8 \cdot 10^{-6}$	$2.0 \cdot 10^{-6}$	$1.7 \cdot 10^{-6}$
Significance Z with $N_{\text{bins}} = 5$, $\text{p}^{\text{syst}} = 50\%$	2	2	2	2	2

Table 12: The expected sensitivity of the HNL signal to the mixing angle $|V_{\mu N}|^2$. The background is modeled using only MC samples. All the signal region requirements listed in Tab. 11 are applied. The top rows show the significance corresponding to the the mixing angles applied in the signal samples which were determined for the specific M_N and $c\tau$ values. The bottom rows show the expected sensitivity of the search when requiring a 2σ significance for the background-only hypothesis. The significance is determined using Eq. 7.2 and binned in the $M(\mu_1, \mu_2, e)$ distribution.

8 Background estimation

The final state signature of the HNL events, $\mu^{\pm}\mu^{\pm}e^{\mp}$, is selected such that the majority of events from background processes are rejected. Remaining SM background events, which mimic the HNL signal, have two identified muons and one identified electron originating from a combination of real and fake leptons. In the analysis below, muons and electrons are referred to as *leptons*, which does not include neutrinos and only τ-leptons which have decayed leptonically to a muon or electron.

In this section, the background estimation for the signal region is examined. First, the terminology used to describe real and fake leptons is presented and the most common fake lepton sources for electrons and muons are presented. The background processes in the signal region are then described which is divided into two main categories: background events with two real muons and one real electron; and background events where one or more of the final state muons and electron is a fake lepton. The method used to estimate the fake lepton background is examined and applied for the signal region leptons. Finally, the background estimation method is applied and compared to data. Since data can not be used in the signal region in order to avoid biasing the analysis and stay blinded, control regions which have a composition similar to the signal region are used.

8.1 Real and fake leptons

Real leptons are leptons which are produced isolated as part of a prompt decay e.g. a decay of a W or Z-boson: [1]:

$$\begin{aligned} W^- &\to e^-\bar{\nu}_e/\mu^-\bar{\nu}_\mu, \\ Z/\gamma^* &\to e^-e^+/\mu^-\mu^+. \end{aligned} \tag{8.1}$$

Electrons and muons which originate from a tau-lepton, that was produced in a prompt decay, are also referred to as real:

$$\begin{aligned} &W^- \to \tau^-\bar{\nu}_\tau \text{ and } Z \to \tau^-\tau^+ \\ &\text{with } \tau^- \to e^-\bar{\nu}_e\nu_\tau/\mu^-\bar{\nu}_\mu\nu_\tau. \end{aligned} \tag{8.2}$$

Due to the produced neutrinos, these events are characterized by having more missing transverse energy than the process in 8.1. Furthermore, the proper life-time of the tau-leptons can result in the electrons or muon being slightly displaced from the primary vertex.

Identified leptons which are *fake* arise from different sources of lepton misidentification. Leptons which origin from pile-up and thus are a part of a different hard-scattering event are also referred to as fake. Fake leptons can either be a different objects, which are wrongly identified, (e.g. a pion) or it can be an actual lepton which is not produced in isolation as part of a prompt decay. Electron fakes can arise from both charged hadrons, photon conversions, or semi-leptonic heavy flavor decays. In photon conversions and semi-leptonic heavy flavor decays there is an actual electron present but it is still denoted as a fake. The fake electrons can for example arise through the following interactions:

$$\begin{aligned} \text{Photon conversion: } & \gamma^* \to e^+e^-, \\ \text{Semi-leptonic heavy flavor decay: } & b \to W^-c \to e^-\bar{\nu}_e c. \end{aligned} \tag{8.3}$$

Due to the large mass of muons compared to electrons, they interact less with matter and they are rarely produced through photon conversion in contrast to the electron. Furthermore, muons are the only particles which leave signals in the muon spectrometer, with the exception of possible hadron shower leakage. Muons can thereby be measured independently in the inner-detector and muon spectrometer and a longer lever arm can be used to determine the curvature of the track. This means that muon fakes are more rare than electron fakes. Almost all fake muons arise from semi-leptonic heavy flavor decays or meson decays-in-flight. The small mass difference between the pion/kaon and the muon allows for interactions where the decay only results in a small deflection of the track and therefore the combined track can mistakenly be reconstructed as coming from a single particle and identified as a muon.

8.2 Background estimation strategy

The background events contributing in the signal region are denoted irreducible or reducible. Irreducible background events have three real leptons in the final state, whereas reducible background events have one or more fake leptons in the final state. The reducible and irreducible background events are treated separately in the background estimation. Reducible background events are modeled using MC. Events which contain fake leptons are not expected to be modeled sufficiently accurate in the MC since the misidentification rate is highly dependent on detector effects. The fake factor method, which is a data-driven procedure, is used to estimate the background from these reducible background events. In order to ensure that the MC contribution is purely from real leptons, it is necessary to apply truth matching of the reconstructed leptons as described in Section 6.3.

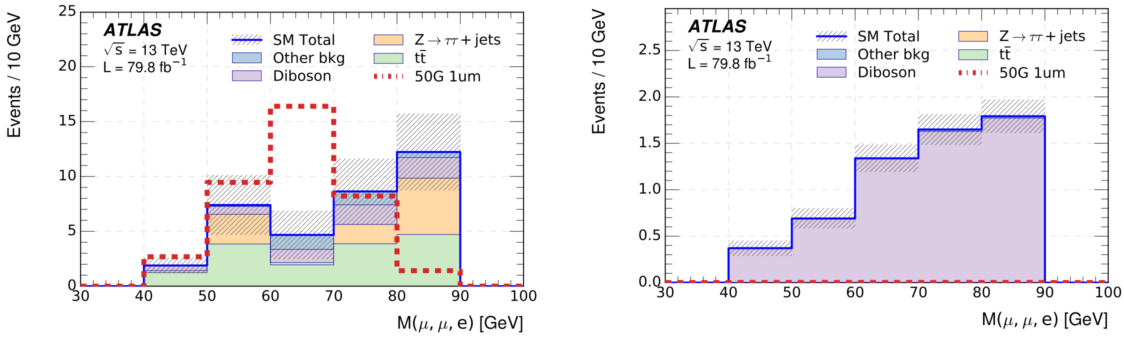
Figure 23 shows the distribution of $M(\mu_1, \mu_2, e)$ in the signal region. On the left hand side, all MC events in the region are shown, whereas on the right hand side, truth matching is applied such that only the irreducible MC background processes with three real leptons are shown. It is found that the majority of the background events contributing to the signal region are reducible background events which contain one or more fake leptons. Thus it is crucial for the analysis that contributions from reducible background processes are estimated accurately. The only non-negligible irreducible background process arise from diboson events. In Table 13, the contributing SM background processes are listed and it is summarised how they arise in the signal region and the estimation strategy applied.

8.2.1 Irreducible background: WZ

The main contribution to the irreducible background comes from the fully leptonic WZ diboson process. This process can mimic the signal region in the decay channel:

$$W^{\pm} \to \mu^{\pm}\nu_{\mu} \text{ and } Z \to \tau^{\pm}\tau^{\mp} \to \mu^{\pm}\nu_{\mu}\nu_{\tau}e^{\mp}\nu_e\nu_{\tau}. \tag{8.4}$$

The τ lepton decay of the Z boson increases the missing transverse energy from neutrinos, so the contribution from this process is reduced by the missing transverse energy requirement: $E_T^{\text{miss}} < 60$ GeV. When the final state includes τ-leptons which decay leptonically, the number of events is also



(a) All MC background processes.

(b) Only irreducible MC background processes.

Figure 23: Distribution of $M(\mu_1, \mu_2, e)$ in the signal region for the signal sample with $M_N = 50$ GeV. All the defined signal region requirements are applied as listed in Tab. 11. The blue line indicates the sum of the SM background processes and the dashed area the corresponding statistical uncertainty. The last bin in the histogram includes all events with values larger than the range of the histogram.

suppressed by the branching ratio since τ-leptons decay hadronically 65% of the time. The following diboson process can also mimic the signal region if one muon has a misidentified charge:

$$W \to e^{\pm}\nu_e \text{ and } Z \to \mu^{\pm}\mu^{\mp}. \tag{8.5}$$

Charge misidentification is very rare for muons since muons can be measured independently in the inner-detector and the muon spectrometer and thereby has a long lever arm to determine curvature of the track.

8.3 Fake background

The fake factor method [62] is a data-driven method developed to estimate the background contribution from processes where one or more leptons are fake leptons. As defined above, the term *fake* leptons refers to both non-prompt leptons and other misidentified objects which are reconstructed as a lepton. In this section, the fake factor method is first presented. The usage of the fake factor method in the analysis is then described. In order to successfully apply the fake factor method, it is crucial to have a good understanding of the background composition and thereby the kinematics of the events which are to be estimated. Composition studies are thus made for each of the three final state leptons. Thereafter, the muon and electron fake factors used to estimate the fake contribution from the subleading muon and electron respectively are studied separately.

8.3.1 The Fake Factor Method

In the fake factor method, the number of background events with one of more fake leptons is estimated using fake factors (F). The fake factors are measured on a dataset which is different from the one used in the signal region, i.e. exactly two leptons are required where the signal region has three leptons. This region is called the *measurement region*. The composition of the fake leptons in the measurement region needs to be similar to the fake leptons in the signal region.

The fake factor is determined using two sets of lepton identification criteria: *ID* criteria which is the signal lepton identification used in the analysis; and *anti-ID* criteria which is designed to be enriched in fake leptons. The anti-ID criteria are defined by relaxing or inverting one or more of the ID criteria applied for the identification and isolation variables.

The fake factor is defined as the ratio between the number of fake leptons in the measurement region which satisfy the ID criteria and anti-ID criteria respectively:

$$F(i) = \frac{N^{\text{data}}_{\text{ID,MR}}(i) - N^{\text{real MC}}_{\text{ID,MR}}(i)}{N^{\text{data}}_{\text{anti-ID,MR}}(i) - N^{\text{real MC}}_{\text{anti-ID,MR}}(i)}, \tag{8.6}$$

Background process	**Origin in signal region**	**Estimation strategy**
	Irreducible SM background processes	
$WZ(\to \tau\tau)$ $(3l)$	Can mimic signal signature	Monte Carlo
$WZ(\to \mu\mu)$ $(3l)$	Charge misidentification of muon	Monte Carlo
ZZ $(4l)$	Missed 4th lepton	Monte Carlo
$t\bar{t}W$ $(3l)$	b-jet fails identification	Monte Carlo
$t\bar{t}Z$ $(4l)$	b-jet fails identification, missed 4th lepton	Monte Carlo
$W^{\pm}W^{\pm}W^{\mp}$ $(3l)$	Can mimic signal signature	Monte Carlo
$W^{\pm}W^{\mp}Z$ $(4l)$	Missed 4th lepton	Monte Carlo
	Reducible SM background processes	
$t\bar{t}$, tW, $t\bar{t}V$ $(2l)$	Jet fakes 3rd lepton	Data-driven*
$Z(\to \mu\mu)$+jets	Jet fakes electron and charge misidentification	Data-driven*
$Z(\to \tau\tau)$+jets	Jet fakes muon	Data-driven*
WW $(2l)$	Jet fakes 3rd lepton	Data-driven*
t, W+jets, $VV(1l)$	Jets fake 2 leptons	Data driven*
Other processes $(0l)$	Jets fake all 3 leptons	Data-driven*

Table 13: Summary of the background processes for a trilepton final state.
*: Except for reducible background events containing a fake leading muon and real subleading muon and electron. These events are estimated using MC.

where N is the number of events, and i refers to the i'th bin of F which is typically in terms of p_T and/or more variables which F depend on. The subscript "MR" indicates that the measurement region is used. By subtracting the number of MC events where the lepton is real, denoted $N^{\text{real MC}}_{\text{ID,MR}}$ and $N^{\text{real MC}}_{\text{anti-ID,MR}}$, it is ensured that both the numerator and denominator consist of only fake lepton events to avoid double counting events. The relevant binning, which should be used for the fake factor, depends on how the MR and the ID and anti-ID criteria are defined. When determining the binning, it is attempted to model the shape of the fake factor distribution as accurate as possible while still maintaining enough statistics to have a low statistical uncertainty on the fake factor.

When estimating the reducible background in the signal region, the anti-ID criteria are again applied to a control region identical to the signal region except that the ID criteria for the lepton is replaced with anti-ID criteria. The estimated number of background events in the application region, which is the signal region or a control region, with a fake lepton is:

$$N^{\text{fake}}_{\text{ID,SR}} = \sum_i \left[N^{\text{data}}_{\text{anti-ID,SR}}(i) - N^{\text{real MC}}_{\text{anti-ID,SR}}(i)\right] F(i), \tag{8.7}$$

where i again refers to the binning of the fake factor. The subscript "SR" indicates that the signal region is used. A schematic presentation of the fake factor regions is shown in Fig. 24.

For a final state with two possibly fake leptons, i.e. a muon and electron, the estimate of the fake background in the signal region is:

$$\begin{aligned}
N^{\text{fake}}_{\text{ID,ID}} =& N^{\text{data}}_{\text{ID,ID}} - N^{\text{real MC}}_{\text{ID,ID}} \\
=& \sum_i \left[N^{\text{data}}_{\text{anti-ID,ID}}(i) - N^{\text{real MC}}_{\text{anti-ID,ID}}(i)\right] F_\mu(i) + \sum_j \left[N^{\text{data}}_{\text{ID,anti-ID}}(j) - N^{\text{real MC}}_{\text{ID,anti-ID}}(j)\right] F_e(j) \\
&- \sum_i \sum_j \left[N^{\text{data}}_{\text{anti-ID,anti-ID}}(i,j) - N^{\text{real MC}}_{\text{anti-ID,anti-ID}}(i,j)\right] F_\mu(i) F_e(j),
\end{aligned} \tag{8.8}$$

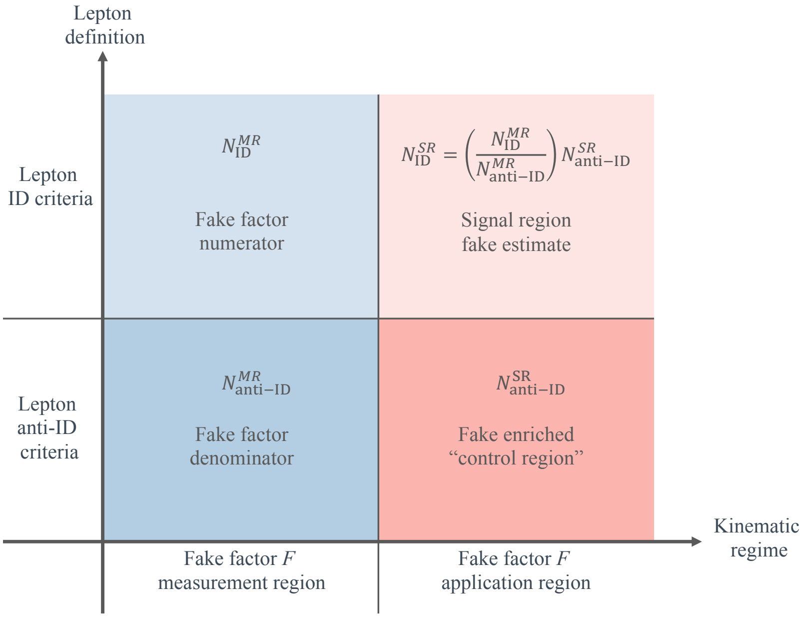


Figure 24: Schematic presentation of the regions applied in the fake factor method. The four regions are all mutually independent.

where F_μ and F_e refer to the fake factor associated to the possibly fake muon and electron in the final state respectively. Note that the term for events where both leptons satisfy the anti-ID criteria has a negative sign opposite to the terms for events with only one lepton that satisfy the anti-ID criteria. Since fake leptons can satisfy the ID criteria, both the estimate $N^{\text{fake}}_{\text{anti-ID,ID}}$ and $N^{\text{fake}}_{\text{ID,anti-ID}}$ can contain events with two fake leptons, so the additional term is necessary to subtract off the double counting. The equation used to estimate the number of background events with fake leptons can easily be generalized to three leptons.

8.3.2 Common Methods for the Muon and Electron Fake Factors

The measurement regions, in which the fake factors are determined, needs to be independent of the signal region and enriched in fake muons/electrons with a similar composition as the fake muons/-electrons in the signal region. In order to determine the fake factors, both a measurement region and ID and anti-ID criteria for the fake lepton have to be defined. In this section, the possible ID and anti-ID criteria for the muon and electron fakes are examined and the measurement regions determined. Ideally, the fake lepton composition should be identical both when comparing the signal and measurement region and when changing between ID and anti-ID criteria. This will be studied in the following sections using truch matching for the muon and electron fakes separately, so in this section only the general strategy for the fake factors is presented.

The ID criteria for the lepton are identical to the criteria applied in the signal region for the subleading muon/electron. Anti-ID criteria are used to enrich the fake contamination by relaxing or inverting some of the requirements. The identification and isolation requirements for both the subleading muon and electron are already very loose, for which the only viable anti-ID definition is to not apply an isolation requirement.

In order to ensure that the fake factor lepton can have a non-isolated baseline and can be measured at low p_T ($p_T \geq 10$ GeV for both the muon and the electron), a *tag and probe* method is used to measure the fake factors. Exactly two leptons are required so the dataset used for the measurement region is independent from the one used for the signal region. The *tag* lepton is required to satisfy single lepton trigger requirements which are listed in Tab. 14. The *probe* muon/electron is used to determine the fake factor and will be referred to as the *fake factor muon/fake factor electron* and labeled $\mu_{\text{FF}}/e_{\text{FF}}$. In order to choose a measurement region which is enriched in fake $\mu_{\text{FF}}/e_{\text{FF}}$, the tag and probe lepton are required to have the same sign charge.

	Single lepton trigger for tag muon	
	Trigger	Offline requirements for μ_{tag}
2015	`HLT_mu20_iloose_L1MU15`	$p_T > 21$ GeV and `Loose` isolation selection.
2016	`HLT_mu26_ivarmedium`	$p_T > 28$ GeV and track-based isolation: $\sum p_T^{\text{trk}}/p_T < 0.07$ in $\Delta R < 0.3$
2017	Same as for the 2016 dataset	Same as for the 2016 dataset
	Single lepton trigger for tag electron	
	Trigger	Offline requirements for e_{tag}
2015	`HLT_e24_lhmedium_L1EM20VH` or	$p_T > 25$ GeV and `MediumLLH` identification selection.
	`HLT_e60_lhmedium` or	$p_T > 61$ GeV and `MediumLLH` identification selection
	`HLT_e120_lhloose`	$p_T(e) > 121$ GeV and `LooseLLH` identification selection.
2016	`HLT_e26_lhtight_nod0_ivarloose` or	$p_T > 27$ GeV, `TightLLH` identification selection and track-based isolation: $\sum p_T^{\text{trk}}/p_T < 0.1$ in $\Delta R < 0.2$
	`HLT_e60_lhmedium_nod0` or	$p_T > 61$ GeV and `MediumLLH` identification selection.
	`HLT_e140_lhloose_nod0`	$p_T > 141$ GeV and `LooseLLH` identification selection.
2017	Same as for the 2016 dataset	Same as for the 2016 dataset

Table 14: The lowest un-prescaled single lepton triggers [46] which were used to collect the dataset during the three data taking periods 2015-2017. They are applied in the measurement regions. When "nod0" is applied in the trigger name, it is indicating that no transverse impact parameter requirements are made. "L1" in the trigger name indicates which Level-1 trigger the event should be seeded by.

Table 15 shows the measurement region definition. In order to ensure a similar composition of fake leptons in the measurement regions as in the signal region, the $E_T^{\text{miss}} < 60$ GeV requirement is also applied in the measurement region. The tag muon is required to be as tight as the leading muon in the signal region in order to ensure similar kinematics. The tag electron is similarly required to fulfill tight identification criteria and isolation selection and the recommended impact parameter requirements.

The impact parameter requirements $|\Delta z_0 \sin(\theta)| < 0.5$ mm and $|d_0|/\sigma(d_0) < 10$ can also be inverted for the anti-ID definition. This is not used since the impact parameter requirements are primarily suppressing pile-up events in the signal region. Since the composition of both fake muons and electrons in the SR is mostly from semi-leptonic heavy flavor decays, it is not desired to make the anti-ID region more pile-up enriched.

The fake factor depends on kinematic variables of the event which can be the lepton transverse momentum, the number of jets (specifically b-jets), the lepton impact parameters, the missing transverse energy and the pseudorapidity of the fake lepton. The dependency of the different variables is related to both the ID and anti-ID criteria used and the composition of the fake lepton in the MR. When determining the binning of the fake factor, it is important to get the right shape of the fake factor while still maintaining a reasonable amount of statistics.

8.3.3 Fake composition studies with MC

The composition of leptons is studied in MC using truth information, see Sec. 6.3. Since the fake factors often depend on the p_T of the lepton, the composition is shown as a function of p_T for the leading muon, subleading muon and electron. When applying studies of the fake composition in the MC, it is important to keep in mind that fake leptons are expected to be poorly modeled in

	Measurement region requirements	
1.	$l^{\pm}_{\text{tag}} l^{\pm}_{\text{probe}}$	
2.	Single lepton trigger matched to l_{tag}, see Tab. 14	
3.	$E_T^{\text{miss}} < 60$ GeV	
	Tag muon	**Tag electron**
4.	$p_T \geq 27$ GeV	$p_T \geq 27$ GeV
5.	$\lvert\eta\rvert < 2.5$	$\lvert\eta\rvert < 2.47$ except $1.37 < \lvert\eta\rvert < 1.52$
6.	Identification `Tight` and `Combined`	Identification `TightLLH`
7.	Isolation `FixedCutTightTrackOnly`	Isolation `FixedCutTight`
8.	$\lvert d_0/\sigma(d_0)\rvert < 3$	$\lvert d_0/\sigma(d_0)\rvert < 5$
9.	$\lvert\Delta z_0 \sin\theta\rvert < 0.5$ mm	$\lvert\Delta z_0 \sin\theta\rvert < 0.5$ mm
	Probe muon (μ_{FF})	**Probe electron** (e_{FF})
10.	$p_T \geq 10$ GeV and $p_T < 50$ GeV	$p_T \geq 10$ GeV and $p_T < 50$ GeV
11.	$\lvert\eta\rvert < 2.5$	$\lvert\eta\rvert < 2.47$ except $1.37 < \lvert\eta\rvert < 1.52$
12.	Identification `Loose` and `Combined`	Identification `LooseAndBlayerLLH`
13.	$\lvert d_0\rvert/\sigma(d_0) < 10$	$\lvert d_0\rvert/\sigma(d_0) < 10$
14.	$\lvert\Delta z_0 \sin\theta\rvert < 0.5$ mm	$\lvert\Delta z_0 \sin\theta\rvert < 0.5$ mm
ID	`FixedCutLoose` isolation	`Gradient` isolation
Anti-ID	NOT `FixedCutLoose` isolation	NOT `Gradient` isolation

Table 15: Measurement region definitions.

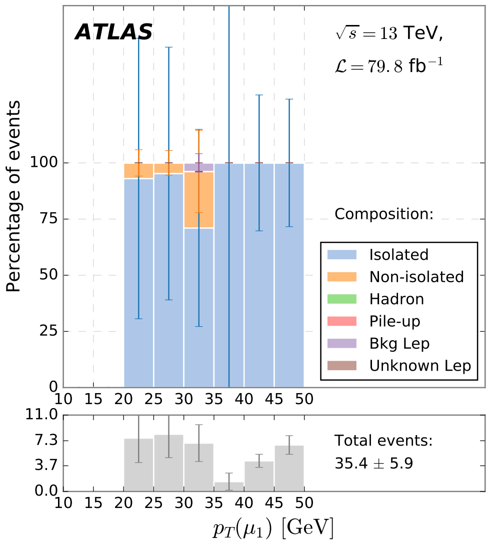
MC which is the reason why it is necessary to apply the data-driven fake factor method. The composition study is used since it provides the best available overview of the topology of the reducible background events but it is not expected to be complete.

The categories used in the composition study are:

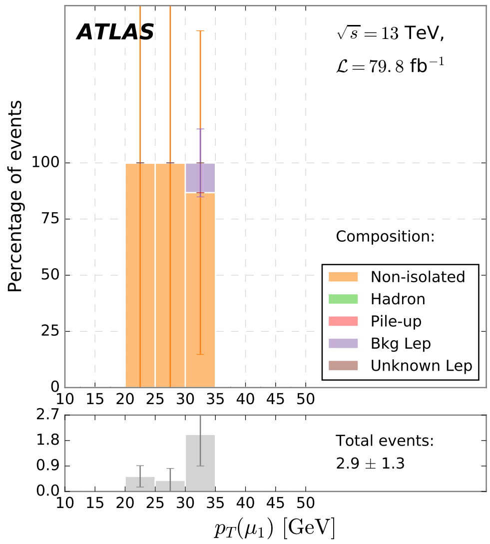
Isolated:	Real leptons.
Non-isolated:	Typically leptons from heavy flavor decays.
Hadron:	Lepton fakes from light flavor hadrons.
Pile-up:	Unknown leptons which usually come from pile-up vertices.
Bkg Lep:	Leptons from e.g. photon conversion and decays of light mesons.
Unknown Lep:	No production vertex found or the algorithm failed to define the origin.

Leptons which are real are in the *isolated* category for truth matching, whereas all the other truth categories describe different fake lepton types. When applying the truth matching algorithm, both a truth type and origin is provided for each lepton. The main categories defined above refer to the truth type of the lepton.

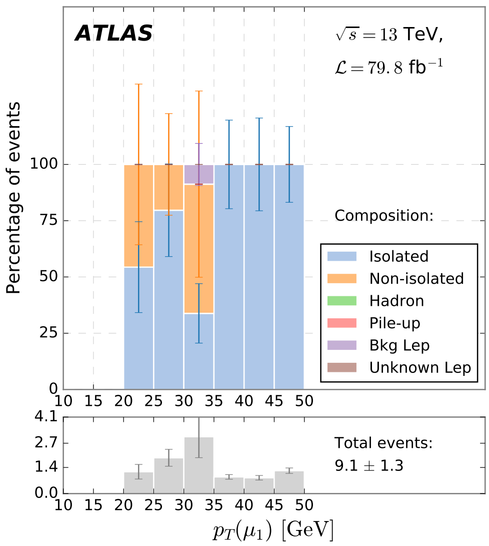
The composition of the leading muon is shown in Fig. 25. As mentioned above, the truth composition is studied as a function of p_T in the range from 10-50 GeV using intervals of 5 GeV. Table 16 shows both the truth matching type and origin for the leading muon. Four different regions are included: the signal region where all leptons are required to satisfy the ID criteria; and the three regions used to estimate the data-driven background, see Eq. 8.8, for which either the subleading muon and/or electron are required to satisfy the anti-ID criteria. It is found that it is not necessary to have a fake factor for the leading muon since it is almost always a real muon.



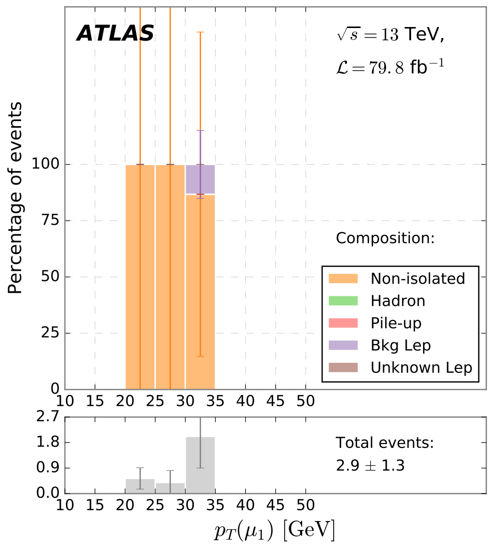
(a) No truth matching requirements.



(b) Fake μ_1.

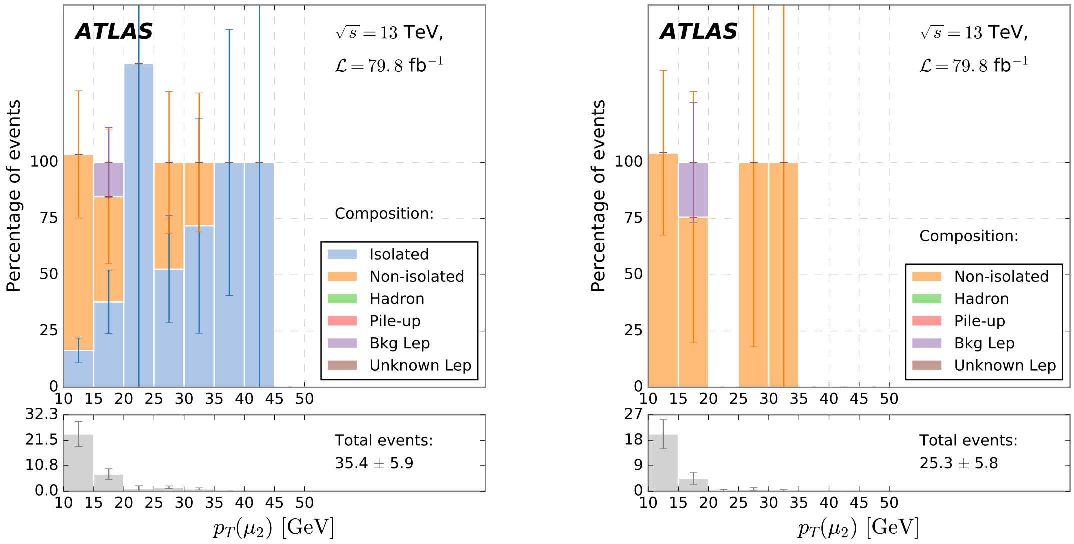


(c) Real μ_2 and real e.
The background contribution in the signal region which is modeled by MC.



(d) Fake μ_1, real μ_2 and real e.
Same composition as in Fig. 25b. Shows that there is no multi-fake events in MC where μ_1 is fake.

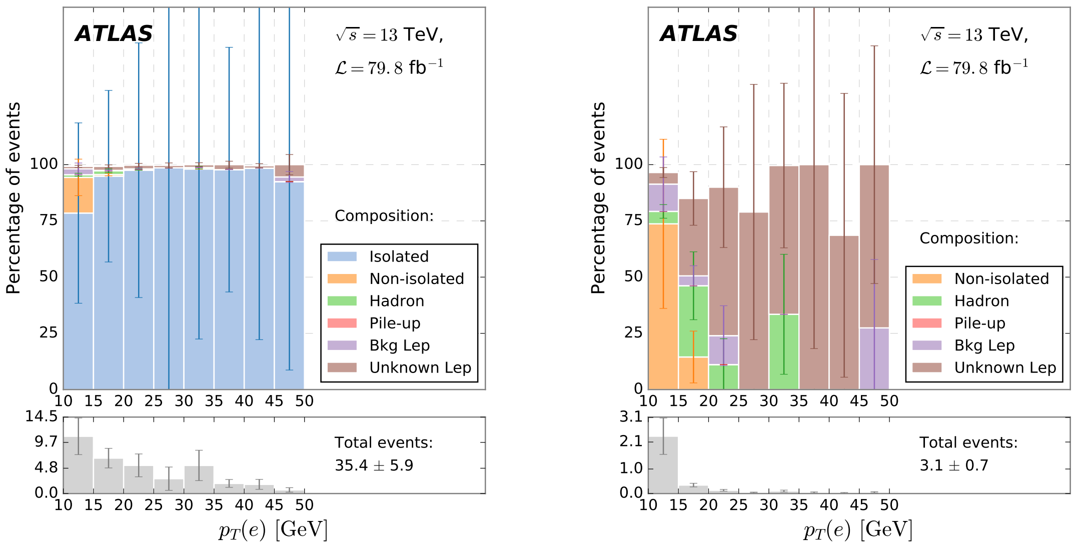
Figure 25: Composition of μ_1 in the signal region as a function of $p_T(\mu_1)$. The final bin contains all the events with $p_T > 45$ GeV. The main plot shows the composition as the percentage of events in the p_T bin. In the subplot, the number of events in each p_T bin is shown as well as the total number of events.



(a) No truthmatching requirements.

(b) Fake μ_2.

Figure 26: Composition of μ_2 in the signal region as a function of $p_T(\mu_2)$. The main plot shows the composition as the percentage of events in the p_T bin. In the subplot, the number of events in each p_T bin is shown as well as the total number of events.



(a) No truthmatching requirements.

(b) Fake e.

Figure 27: Composition of e in the signal region as a function of $p_T(e)$. The final bin contains all the events with $p_T > 45$ GeV. The main plot shows the composition as the percentage of events in the p_T bin. In the subplot, the number of events in each p_T bin is shown as well as the total number of events.

Type	Origin	Signal region All ID	Anti-ID μ_2	Anti-ID e	Anti-ID μ_2 and e
IsoMuon	NonDefined	0.6±0.1	–	0.1±0.0	–
IsoMuon	Top	14±2	43±4	4.8±1.1	15±2
IsoMuon	WBoson	6.3±0.9	1±20	12±6	47±25
IsoMuon	ZBoson	12±5	6±7	5±3	2±2
NonIsoMuon	CharmedMeson	1±1	0.7±0.7	–	–
NonIsoMuon	BottomMeson	1.9±0.7	–	0.2±0.2	–
BkgMuon	CCBarMeson	0.3±0.3	–	–	–
NonIsoMuon	TauLep	–	0.1±0.1	–	–
Number of events with real μ_1:		33±6	51±21	21±7	64±25
Number of events with fake μ_1:		3±1	0.8±0.7	0.2±0.2	–
Total number of events:		35±6	52±21	21±7	64±25

Table 16: Composition of the leading muon in the signal region classified using truth matching. The ID and anti-ID criteria are shown in Tab. 15. Negative number of events can appear due to negative MC weights, see Sec. 6.2. The applied MC events are scaled to an integrated luminosity of $\mathcal{L} = 79.8$ fb^{-1}.

Type	Origin	Signal region All ID	Anti-ID μ_2	Anti-ID e	Anti-ID μ_2 and e
Unknown	NonDefined	-1±1	2±1	–	0.3±0.3
IsoMuon	NonDefined	0.6±0.1	–	0.1±0.0	–
IsoMuon	Top	2.1±0.8	–	0.2±0.2	–
IsoMuon	WBoson	1.3±0.2	–	0.2±0.1	–
IsoMuon	ZBoson	6±1	1.2±0.7	0.7±0.1	–
IsoMuon	Higgs	–	–	–	–
NonIsoMuon	TauLep	0.3±0.3	0.9±0.5	–	0.4±0.3
NonIsoMuon	CharmedMeson	4±3	-16±19	3±3	13±11
NonIsoMuon	BottomMeson	17±4	60±8	16±6	27±9
NonIsoMuon	CharmedBaryon	3±3	0.3±0.3	–	–
NonIsoMuon	BottomBaryon	1.5±1.0	4±3	1.1±0.5	23±21
BkgMuon	PhotonConv	–	0.3±0.3	–	–
BkgMuon	CCBarMeson	1±1	–	–	0.3±0.3
Number of events with real μ_2:		10±1	1.2±0.7	1.3±0.2	–
Number of events with fake μ_2:		25±6	51±21	20±7	65±26
Total number of events:		35±6	52±21	21±7	64±25

Table 17: Composition of the subleading muon in the signal region classified using truth matching. The ID and anti-ID criteria are shown in Tab. 15. Negative number of events can appear due to negative MC weights, see Sec. 6.2. The applied MC events are scaled to an integrated luminosity of $\mathcal{L} = 79.8$ fb^{-1}.

Figure 26 shows the composition of the subleading muon and Tab. 17 shows the truth origin and type for the subleading muon. The majority of the background events have a fake subleading muon which primarily arise from semi-leptonic decays of b-quarks in mesons.

Figure 27 shows the composition of the electron. Due to the very low statistics for the electron fake contributions, the uncertainties on the composition components are very large. The truth origin and type for the electron are shown in Tab. 18. Table 18 shows a detailed list of the truth composition for the electron. It is found that the primary source of fake electrons is again through semi-leptonic decays of b-quarks in mesons but there is a contribution from both misidentified light flavor jets and leptons from e.g. photon conversion. This means that, as expected, the muon and electron fake factors will not necessarily be equal.

Type	Origin	Signal region All ID	Anti-ID μ_2	Anti-ID e	Anti-ID μ_2 and e
UnknownElectron	NonDefined	0.5±0.1	–	0.1±0.0	–
IsoElectron	Top	15±2	38±3	1.4±0.6	4±1
IsoElectron	WBoson	1.8±0.7	6±1	0.3±0.2	0.7±0.4
IsoElectron	ZBoson	16±6	4±6	3±3	0.8±2
IsoElectron	Higgs	–	–	–	–
NonIsoElectron	TauLep	0.1±0.0	0.5±0.4	0.6±0.4	0.8±0.5
NonIsoElectron	CharmedMeson	–	1.1±0.6	0.1±0.0	0.8±0.5
NonIsoElectron	BottomMeson	1.4±0.6	11±5	10±5	38±22
NonIsoElectron	BottomBaryon	0.3±0.3	4±4	0.1±0.1	3±4
BkgElectron	PhotonConv	0.4±0.3	3±2	0.4±0.3	0.3±0.3
IsoMuon	ZBoson	0.1±0.0	–	1±1	1±1
Hadron	NonDefined	0.3±0.1	-16±19	4±2	13±11
NonIsoElectron	CharmedBaryon	–	–	–	0.3±0.3
Number of events with real e:		32±6	48±7	5±3	6±2
Number of events with fake e:		3±1	3±20	16±6	57±25
Total number of events:		35±6	52±21	21±7	64±25

Table 18: Composition of the electron in the signal region classified using truth matching. The ID and anti-ID criteria are shown in Tab. 15. Negative number of events can appear due to negative MC weights, see Sec. 6.2. The applied MC events are scaled to an integrated luminosity of $\mathcal{L} = 79.8$ fb^{-1}.

8.3.4 Model stability checks

The fake factor model is dependent on the definition of ID and anti-ID criteria which are used to separate real and fake lepton contributions. When the ID and anti-ID criteria are changed, the statistics in both the application region (signal region or control region) and the measurement region are changed accordingly. In order to check the stability of the fake factor method against changes in the ID and anti-ID definition, additional muon and electron fake factors are determined using looser isolation selections than those chosen in the signal optimization. When changing the isolation selection to a looser selection, i.e. more efficient, some of the events which satisfied the anti-ID criteria for the tighter isolation selection will satisfy the ID criteria instead given the looser isolation selection. This results in larger fake factors, see 8.6. These regions will be used to test the model stability and assign a systematic uncertainty in Sec. 9.

The tested isolation selections are all listed in Tab. 19. Only the isolation selection is changed, so all additional requirements used to define the muon- and electron measurement regions will remain constant. In Tab. 20, the ID and anti-ID criteria used in all the applied fake factor measurement regions are listed, which shows that they all are defined using a non-isolated baseline.

Working point	Object	Calorimeter isolation	Track isolation
`Loose`	$e + \mu$	$\epsilon = 99\%$	$\epsilon = 99\%$
`Gradient`	$e + \mu$	$\epsilon =$(0.1143*pT[GeV]+92.14)%	$\epsilon =$(0.1143*pT[GeV]+92.14)%
`FixedCutLoose`	μ	Cut: topoetcone20/pT<0.3	Cut: ptvarcone30/pT<0.15
`Loosened FixedCutLoose`	μ	Cut: topoetcone20/pT<0.4	Cut: ptvarcone30/pT<0.20

Table 19: Definitions of the isolation selections applied to define the ID and anti-ID criteria for the muon and electron fake factors. The combined isolation efficiency for the `Loose` WP is 99%. For the `Gradient` WP, the combined isolation efficiency is $\epsilon = 90\%/99\%$ for $p_T = 25/60$ GeV.

Fake factor	ID criteria	Anti-ID criteria
Muon	Satisfy `FixedCutLoose`	Not satisfy `FixedCutLoose`
Muon	Satisfy `Loosened FixedCutLoose`	Not satisfy `Loosened FixedCutLoose`
Muon	Satisfy `Loose`	Not satisfy `Loose`
Electron	Satisfy `Gradient`	Not satisfy `Gradient`
Electron	Satisfy `Loose`	Not satisfy `Loose`

Table 20: ID and anti-ID criteria used for all the applied fake factor measurement regions.

8.3.5 Muon Fake Factor

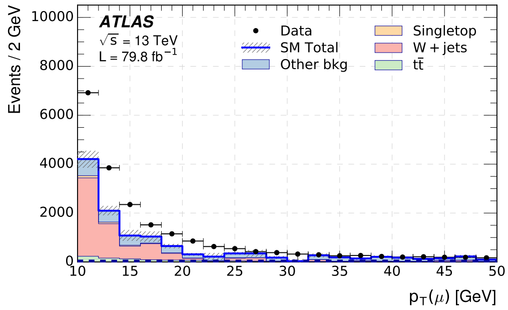
The muon fake factor will be used to estimate the contribution from fake subleading muons in the signal region. In this section, three sets of fake factors will be determined for which the only difference is the isolation selection used for μ_{FF} which define the ID and anti-ID criteria. As mentioned in the previous section, the three isolation selections are: `FixedCutLoose`, `Loosened FixedCutLoose`, and `Loose`, see Tab. 19. In the analysis below, I will focus on the fake factor using the `FixedCutLoose` isolation, which is the isolation selection chosen in the optimization study of the signal region.

The muon measurement region is defined in Tab. 15 where the tag lepton is either an electron or muon and the probe lepton is a muon: $e^{\pm}\mu^{\pm}$ and $\mu^{\pm}\mu^{\pm}$. Due to the very low rate of charge misidentification for muons, it is possible to use the $\mu^{\pm}\mu^{\pm}$ channel without any need to apply a specific cut to reject real muons from Z-boson decays.

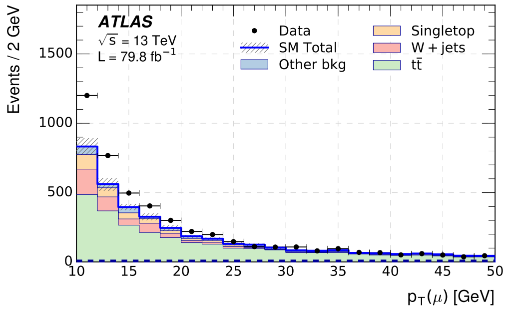
The muon measurement region is primarily targeting events from W+jets and $t\bar{t}$ events. Figure 28 shows the distribution of $p_T(\mu_{\text{FF}})$ in the MR using the ID criteria. The plots show that the MC is similar for the two channels: $e^{\pm}\mu^{\pm}$ and $\mu^{\pm}\mu^{\pm}$, and that the dominating MC is W+jets when a b-jet veto is applied and $t\bar{t}$ dominates when $N_{\text{b-jets}} \geq 1$. It is also shown that either the MC simulations model the data very poorly in this fake-enriched MR or there can be some additional missing processes for which MC does not exist. Similarly, Fig. 29 shows the distribution of $p_T(\mu_{\text{FF}})$ in the MR using the anti-ID criteria.

Figure 30 shows the composition of the μ_{FF} in the muon MR as a function of $p_T(\mu_{\text{FF}})$. The plots show that the majority of μ_{FF} has low p_T from 10-20 GeV and arise from semi-leptonic heavy flavor decays, "non-isolated" type, which corresponds to the composition of the subleading muon in the signal region. When applying a b-jet veto, a part of the fake muons in the MC events are labeled as "unknown" type with "non defined" origin which often refer to fake leptons from pile-up. Assuming that μ_{FF} in these events are indeed originating from pile-up, then it is important that the binning of the fake factor take into account the kinematic differences between fake muons arising from heavy flavor decays and from pile-up.

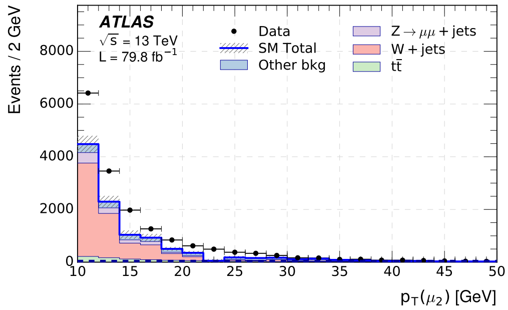
When determining the fake factor, see Eq. 8.6, the number of MC events for which μ_{FF} is real is subtracted from the number of data events. Due to the same-sign requirement for the tag and probe lepton, the amount of MC with a real μ_{FF} is low. The majority of the subtracted MC origins from diboson events for $N_{\text{b-jets}}$=0 and $t\bar{t}$ events for $N_{\text{b-jets}} \geq 1$. The tight requirements applied for the tag lepton, which is the leading muon in the $\mu^{\pm}\mu^{\pm}$ channel and the electron in the $e^{\pm}\mu^{\pm}$ channel, is ensuring that the large majority of the tag leptons are real leptons. But only the fake factor muon is required to be real in the subtracted MC, so the events where the tag lepton is not real can contribute as well. Diboson events from WZ all-leptonic decays can contribute if a lepton from the Z-boson decay is lost. Furthermore, events where one lepton charge is mis-identified can contribute e.g. through Z-boson decay to $\mu^{\pm}\mu^{\mp}$ and $\tau^{\pm}\tau^{\mp}$ and through WW events. Events from $t\bar{t}$ production can also contribute if one of the leptons is charge-flipped or if the tag lepton is originating from an energetic semi-leptonic heavy flavor decay.



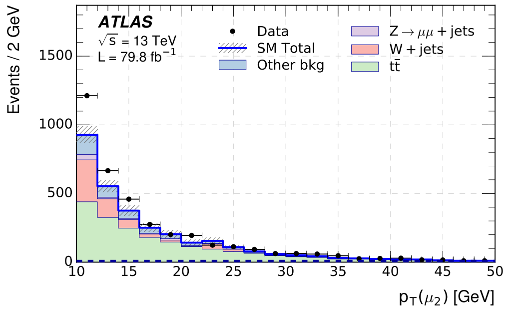
(a) Muon MR channel: $e^{\pm}\mu^{\pm}$. ID criteria. $N_{\text{b-jets}}$=0.



(b) Muon MR channel: $e^{\pm}\mu^{\pm}$. ID criteria. $N_{\text{b-jets}} \geq 1$.

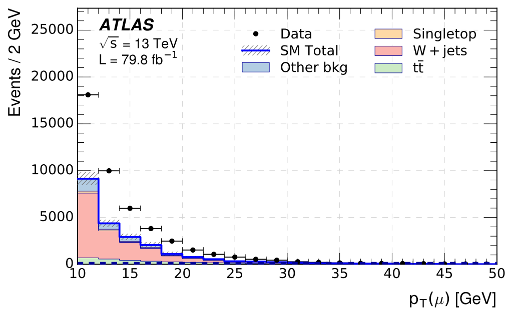


(c) Muon MR channel: $\mu^{\pm}\mu^{\pm}$. ID criteria. $N_{\text{b-jets}}$=0.

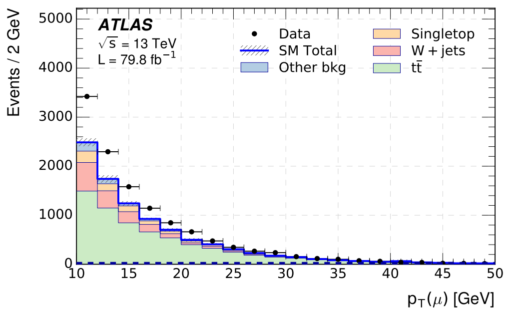


(d) Muon MR channel: $\mu^{\pm}\mu^{\pm}$. ID criteria. $N_{\text{b-jets}} \geq 1$.

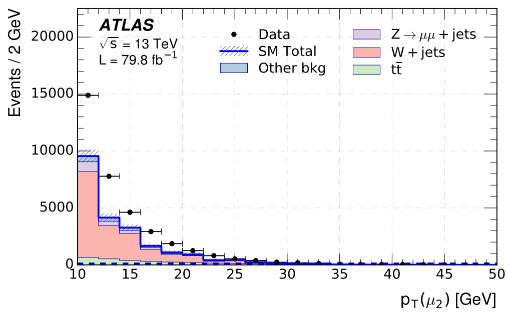
Figure 28: The muon measurement region with ID criteria for μ_{FF}. Distribution of $p_T(\mu_{\text{FF}})$. The blue line indicates the sum of the SM background processes and the dashed area the corresponding statistical uncertainty. The last bin in the histogram includes all events with values larger than the range of the histogram.



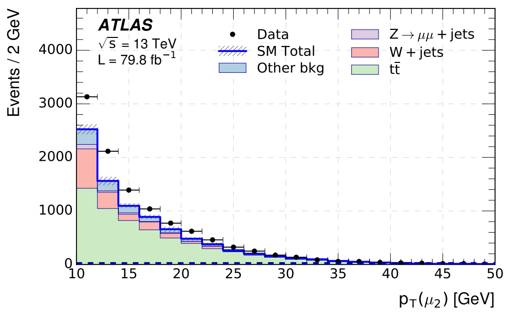
(a) Muon MR channel: $e^{\pm}\mu^{\pm}$. Anti-ID criteria. $N_{\text{b-jets}}$=0.



(b) Muon MR channel: $e^{\pm}\mu^{\pm}$. Anti-ID criteria. $N_{\text{b-jets}} \geq 1$.

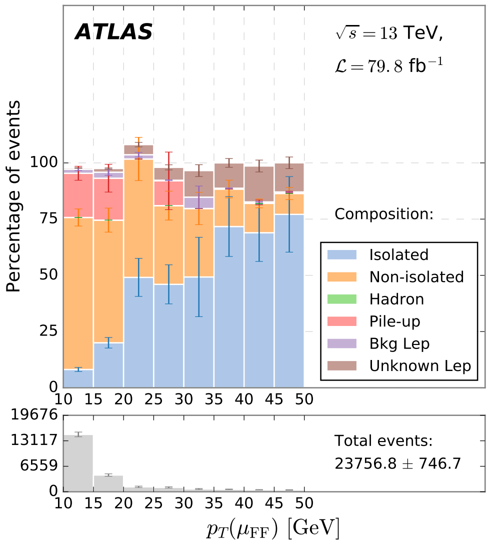


(c) Muon MR channel: $\mu^{\pm}\mu^{\pm}$. Anti-ID criteria. $N_{\text{b-jets}}$=0.

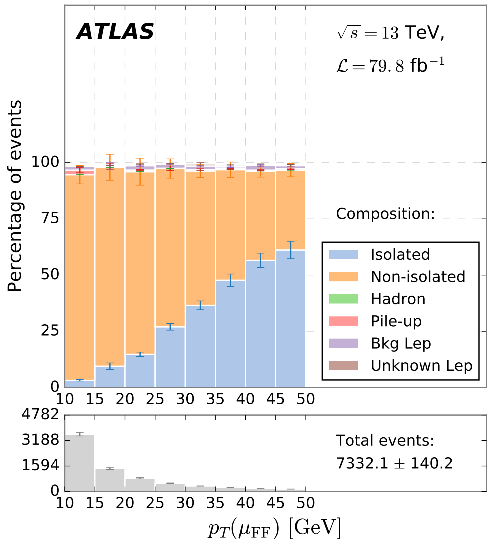


(d) Muon MR channel: $\mu^{\pm}\mu^{\pm}$. Anti-ID criteria. $N_{\text{b-jets}} \geq 1$.

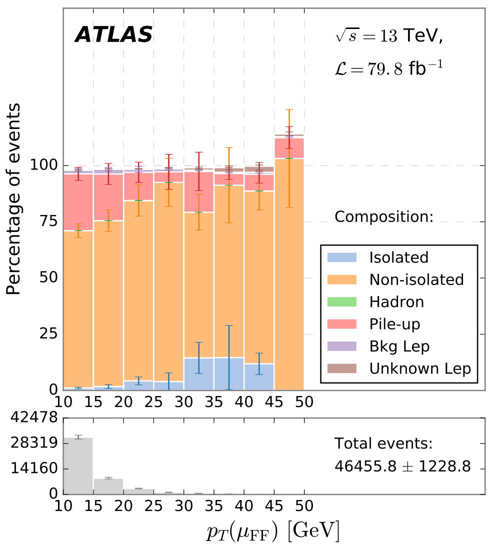
Figure 29: The muon measurement region with anti-ID criteria for μ_{FF}. Distribution of $p_T(\mu_{\text{FF}})$. The blue line indicates the sum of the SM background processes and the dashed area the corresponding statistical uncertainty. The last bin in the histogram includes all events with values larger than the range of the histogram.



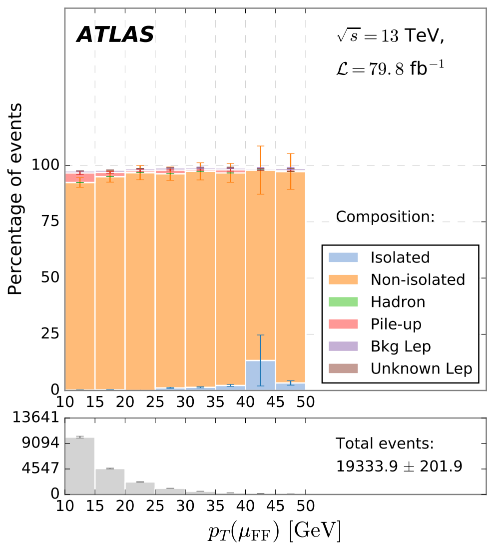
(a) ID criteria for μ_{FF}. $N_{\text{b-jets}}$=0.



(b) ID criteria for μ_{FF}. $N_{\text{b-jets}} \geq 1$.



(c) Anti-ID criteria for μ_{FF}. $N_{\text{b-jets}}$=0.



(d) Anti-ID criteria for μ_{FF}. $N_{\text{b-jets}} \geq 1$.

Figure 30: Composition of μ_{FF} in the muon MR as a function of $p_T(\mu_{\mathrm{FF}})$ when requiring μ_{FF} to satisfy the ID criteria (top plots) and the anti-ID criteria (bottom plots). The final bin contains all the events with $p_T > 45$ GeV. The main plots show the composition as the percentage of events in the p_T bin, whereas the subplots show the number of events in each p_T bin and the total number of events.

Table 21 provides an overview of the number of data events, number of MC events with a real μ_{FF} and the average fake factor in each of the muon MRs for the three considered isolation WPs. It is shown that the `Loose` isolation selection is with an average fake factor of 5.8 very bad at discriminating fake and real leptons compared to the `FixedCutLoose` and `Loosened FixedCutLoose` isolation selections. In the next section, the muon fake factor is determined for the case where the ID and anti-ID criteria are defined using the `FixedCutLoose` isolation selection.

	$N_{\mathrm{ID}}^{\mathrm{data}}/10^3$	$N_{\mathrm{anti\text{-}ID}}^{\mathrm{data}}/10^3$	$N_{\mathrm{ID}}^{\mathrm{real\ MC}}/10^3$	$N_{\mathrm{anti\text{-}ID}}^{\mathrm{real\ MC}}/10^3$	F_{average}
`FixedCutLoose`	45.4±0.2	102.7±0.3	4.7±0.14	0.39±0.05	0.35
`Loosened FixedCutLoose`	56.7±0.2	91.5±302	4.8±0.15	0.27±0.03	0.57
`Loose`	126.8±0.4	21.4±0.15	4.9± 0.14	0.21±0.03	5.8

Table 21: Number of data events, number of MC events (with a real μ_{FF}) and the average fake factor in the applied muon measurement regions. The isolation selections used to determine the ID and anti-ID criteria are listed.

8.3.6 Binning of the Muon Fake Factor

In this section, the binning applied for the muon fake factor is determined. The isolation selection `FixedCutLoose` is used for μ_{FF}. The fake factor dependency is determined for the following variables: $N_{\text{b-jets}}$, $p_T(\mu_{\mathrm{FF}})$, $|\eta|(\mu_{\mathrm{FF}})$, $|d_0|/\sigma(d_0)(\mu_{\mathrm{FF}})$, $|\Delta z_0 \sin(\theta)|(\mu_{\mathrm{FF}})$, N_{jets}, and E_T^{miss}. In order to avoid large errors on the fake factors while still maintaining a good description of the shape, two fake factor bins are merged if they are consistent within their statistical uncertainties. The size of the considered muon fake factor appear to depend on the variables: $N_{\text{b-jets}}$, $p_T(\mu_{\mathrm{FF}})$, N_{jets}, and $|d_0|/\sigma(d_0)(\mu_{\mathrm{FF}})$. The fake factor bins are determined in the following order to ensure that correlations and statistics are taken into account properly:

1. Binning as a function of $N_{\text{b-jets}}$.

1. Binning each $N_{\text{b-jets}}$ bin as a function of $p_T(\mu_{\mathrm{FF}})$.

2. Binning each $N_{\text{b-jets}}$ and $p_T(\mu_{\mathrm{FF}})$ bin as a function of N_{jets}.

3. Binning each $N_{\text{b-jets}}$, $p_T(\mu_{\mathrm{FF}})$ and N_{jets} bin as a function of $|d_0|/\sigma(d_0)(\mu_{\mathrm{FF}})$.

Statistically consistent fake factor bins will be merged doing each step. In all plots of the fake factors in this section, the range of each bin is shown using horizontal errorbars. Statistical uncertainties are shown on the vertical errorbars.

First, the fake factor dependency on $N_{\text{b-jets}}$ is considered. Figure 31 shows the fake factor as a function of $N_{\text{b-jets}}$. The fake factor clearly depends on $N_{\text{b-jets}}$, so the fake factor is divided into the two regions: $N_{\text{b-jets}} = 0$ and $N_{\text{b-jets}} \geq 1$. Since a b-jet veto is applied in the signal region, the fake factors determined in the $N_{\text{b-jets}} = 0$ measurement region will be used to determine the data-driven fake background in the signal region. Still, fake factors for $N_{\text{b-jets}} \geq 1$ will be constructed too in order to check the stability of the model and estimate the fake background in a control region where $N_{\text{b-jets}} \geq 1$. In the following section, I will focus on the fake factor for which $N_{\text{b-jets}} = 0$.

Figure 32 shows the fake factor as a function of the six variables: $p_T(\mu_{\mathrm{FF}})$, $|\eta|(\mu_{\mathrm{FF}})$, $|d_0|/\sigma(d_0)(\mu_{\mathrm{FF}})$, $|\Delta z_0 \sin(\theta)|(\mu_{\mathrm{FF}})$, N_{jets}, and E_T^{miss}. It is found that the fake factor shape is most dependent on $p_T(\mu_{\mathrm{FF}})$ and this variable will be applied to bin the fake factor. Figure 102, included in appendix Sec. F, shows in detail the fake factor dependency on the transverse momentum of the fake factor muon and the applied binning of the fake factor as a function of $p_T(\mu_{\mathrm{FF}})$. The bins are defined in the following ranges of $p_T(\mu_{\mathrm{FF}})$: 10-20 GeV, 20-30 GeV, 30-40 GeV, and 40-50 GeV.

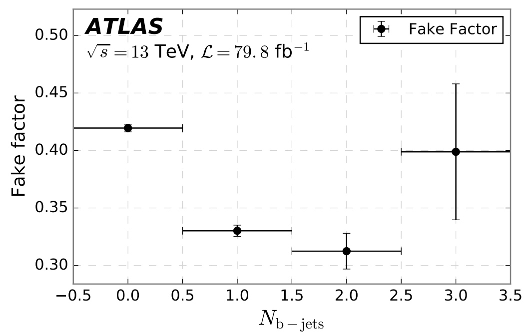
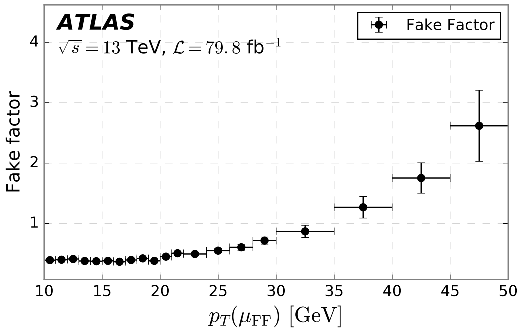


Figure 31: Muon fake factor as a function of $N_{\text{b-jets}}$. The ID and anti-ID criteria are defined using the `FixedCutLoose` isolation selection.

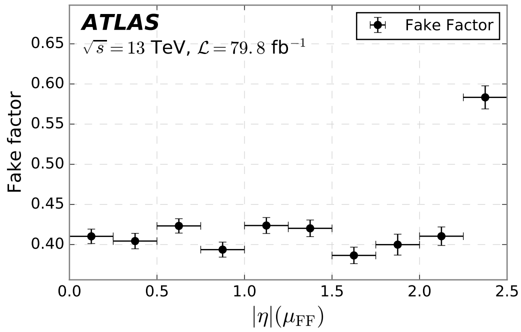
The fake factor dependency on the variables $|d_0|/\sigma(d_0)(\mu_{\mathrm{FF}})$, E_T^{miss}, and N_{jets} is examined again for the four fake factor bins in $p_T(\mu_{\mathrm{FF}})$. The figures applied in this study are all included in appendix Sec. F. Figure 103 shows the fake factor as a function of E_T^{miss}. There is no clear tendency in the E_T^{miss} dependency across the four $p_T(\mu_{\mathrm{FF}})$ bins so this variables will not be used to bin the fake factor. Figure 104 shows the fake factor as a function of $|d_0|/\sigma(d_0)(\mu_{\mathrm{FF}})$. The plots show that the fake factor varies as a function of $|d_0|/\sigma(d_0)(\mu_{\mathrm{FF}})$ and there is a clear tendency that the fake factor decreases for an increasing $|d_0|/\sigma(d_0)(\mu_{\mathrm{FF}})$. 105 shows the fake factor as a function of N_{jets}. It is found that the fake factors decrease as a function of N_{jets} for all the four p_T bins. It is chosen to use the variable N_{jets} to bin the fake factors and consider the fake factor dependency of $|d_0|/\sigma(d_0)(\mu_{\mathrm{FF}})$ for each of the fake factor bins in $p_T(\mu_{\mathrm{FF}})$ and N_{jets}. Two bins are used as a function of N_{jets} defined by: $N_{\mathrm{jets}} \leq 2$ and $N_{\mathrm{jets}} > 2$. In Fig. 105, the applied bins in N_{jets} are shown on the figures in the right hand side.

Figure 106 shows the fake factor as a function of $|d_0|/\sigma(d_0)(\mu_{\mathrm{FF}})$ for the four p_T bins with $N_{\mathrm{jets}} \leq 2$. Similarly, Fig. 107 shows the fake factor as a function of $|d_0|/\sigma(d_0)(\mu_{\mathrm{FF}})$ for the four p_T bins with $N_{\mathrm{jets}} > 2$. Again, it is found that the fake factor decreases when $|d_0|/\sigma(d_0)(\mu_{\mathrm{FF}})$ increase. Two bins are applied as a function of $|d_0|/\sigma(d_0)(\mu_{\mathrm{FF}})$: $|d_0|/\sigma(d_0)(\mu_{\mathrm{FF}}) \leq 2$ and $2 < |d_0|/\sigma(d_0)(\mu_{\mathrm{FF}}) < 10$. The final muon fake factors are shown in Fig. 33. For the fake factor shown in Fig. 33a with the binning 10 GeV$p_T(\mu_{\mathrm{FF}}) < 20$ GeV and $N_{\mathrm{jets}} > 2$, the fake factor estimated for the two bins in $|d_0|/\sigma(d_0)(\mu_{\mathrm{FF}})$ are consistent within the statistical uncertainties so the two bins are merged.

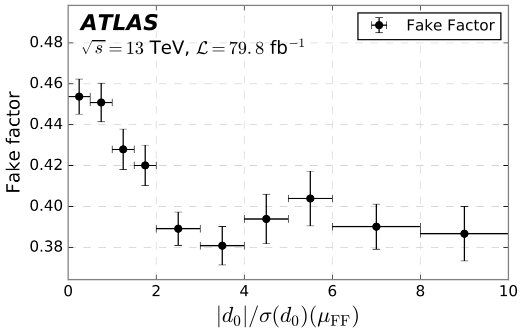
A study of the fake factor dependencies is similarly made for the muon measurement region for which $N_{\text{b-jets}} \geq 1$ and the measurement regions applying looser isolation selection. These results will be used later to check closure and solidity of the method against data. Studies showing the fake factor dependencies in these measurement regions are included in appendix Sec. F. Table 22 shows the applied binning of the muon fake factors in all the considered measurement regions defined by the isolation selection of μ_{FF} and the $N_{\text{b-jets}}$ requirement. All additional requirements for the measurement regions are constant, as listed in Tab. **??**.



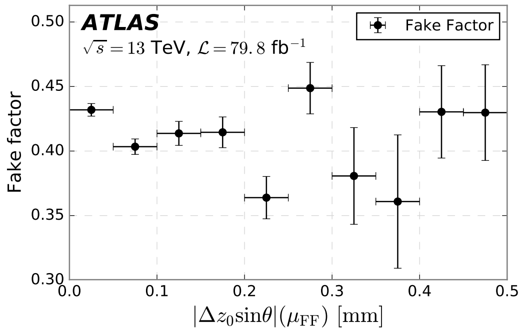
(a) Fake factor as a function of $p_T(\mu_{\rm FF})$.



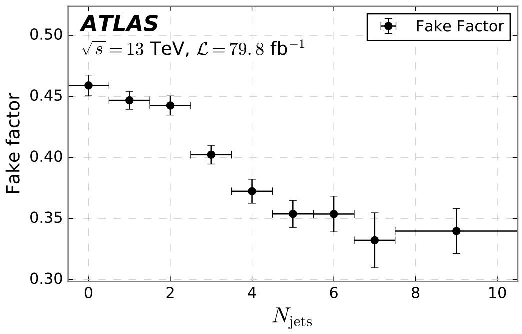
(b) Fake factor as a function of $|\eta|(\mu_{\rm FF})$.



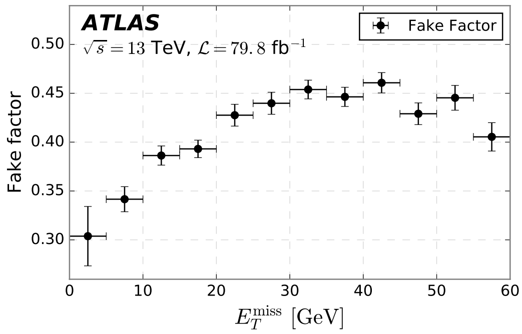
(c) Fake factor as a function of $|d_0|/\sigma(d_0)(\mu_{\rm FF})$.



(d) Fake factor as a function of $|\Delta z_0 \sin(\theta)|(\mu_{\rm FF})$.

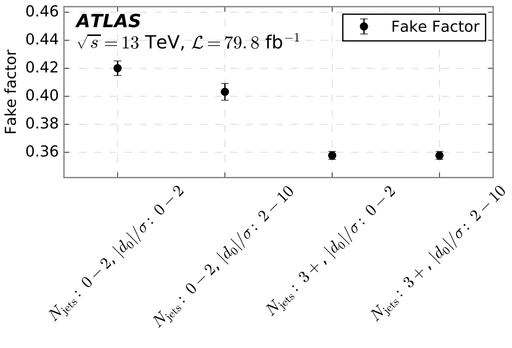


(e) Fake factor as a function of $N_{\rm jets}$.

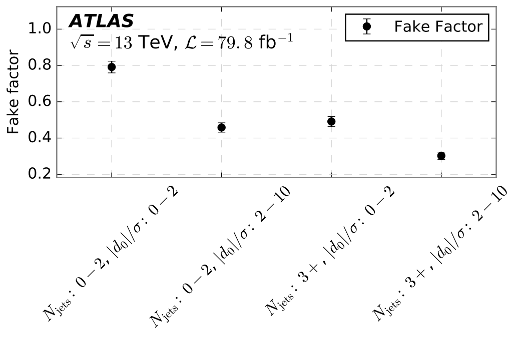


(f) Fake factor as a function of $E_T^{\rm miss}$.

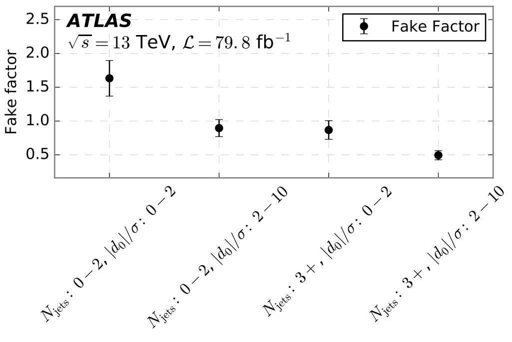
Figure 32: Muon fake factor as a function of: $p_T(\mu_{\rm FF})$; $|\eta|(\mu_{\rm FF})$; $|d_0|/\sigma(d_0)(\mu_{\rm FF})$; $|\Delta z_0 \sin(\theta)|(\mu_{\rm FF})$; $N_{\rm jets}$; and $E_T^{\rm miss}$. The ID and anti-ID criteria are defined using the `FixedCutLoose` isolation in the muon measurement region with $N_{\text{b-jets}} = 0$.



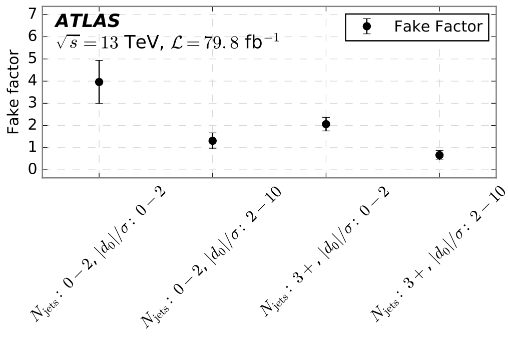
(a) Bin: 10 GeV $\leq p_T(\mu_{\mathrm{FF}}) < 20$ GeV



(b) Bin: 20 GeV $\leq p_T(\mu_{\mathrm{FF}}) < 30$ GeV



(c) Bin: 30 GeV $\leq p_T(\mu_{\mathrm{FF}}) < 40$ GeV



(d) Bin: 40 GeV $\leq p_T(\mu_{\mathrm{FF}}) < 50$ GeV

Figure 33: The final muon fake factor binning. The ID and anti-ID criteria are defined using the `FixedCutLoose` isolation in the muon measurement region with $N_{\text{b-jets}} = 0$.

Measurement region		**Applied fake factor bins**			
Isolation	$N_{\text{b-jets}}$	$p_T(\mu_{\mathrm{FF}})$	N_{jets}	$\lvert d_0\rvert/\sigma(d_0)(\mu_{\mathrm{FF}})$	$\lvert\eta\rvert(\mu_{\mathrm{FF}})$
`FixedCutLoose`	0	4 bins	2 bins	2 bins	–
`FixedCutLoose`	≥ 1	3 bins	2 bins	–	–
`Loosened FixedCutLoose`	0	4 bins	2 bins	2 bins	–
`Loosened FixedCutLoose`	≥ 1	4 bins	2 bins	–	–
`Loose`	0	5 bins	2 bins	–	3 bins
`Loose`	≥ 1	4 bins	–	–	3 bins

Table 22: The applied fake factor bins in each of the muon measurement regions. The measurement regions are defined by the isolation selection used for μ_{FF} which defines the ID and anti-ID criteria, see Tab. 20, and the requirement for $N_{\text{b-jets}}$. All the final fake factor bins are included in appendix Sec. F.

8.3.7 Electron Fake Factor

The method used to determine the electron fake factors follows the one used for the muon fake factors in the previous section. For the electron, two sets of fake factors are determined for which the ID and anti-ID criteria are defined using the `Gradient` and `Loose` isolation selections, see Tab. 19. The isolation selection `Loose` is also applied in order to test the model stability for changing ID and anti-ID criteria, see Sec. 8.3.4. Again, the focus in this section will be on the fake factor for the isolation selection chosen in the optimization study of the signal region which is the `Gradient` isolation.

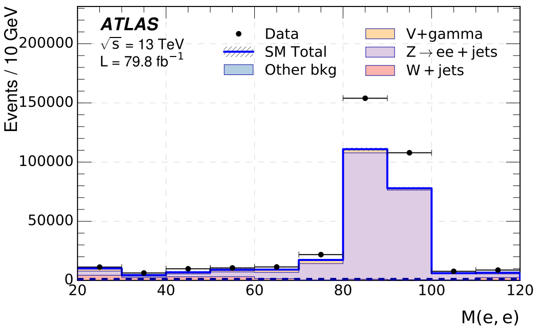
The electron measurement region is defined in Tab. 15 where the tag lepton can be either an electron or a muon: $e^{\pm}e^{\pm}\mu^{\pm}e^{\pm}$. For the channel $e^{\pm}e^{\pm}$, real SM background from the Drell-Yan process $Z \rightarrow e^{\pm}e^{\mp}$ can make a relatively large contribution to the MR due to the higher rate of electrons reconstructed with an opposite charge due to photon conversion. The two electron measurement region channels are studied in Fig. 34 which shows the dilepton invariant mass distribution: $M(e_1, e_2)$ for the channel $e^{\pm}e^{\pm}$ and $M(\mu, e)$ for the channel $\mu^{\pm}e^{\pm}$. For the figures on the right hand side, e_{FF} is required to be real which is exactly the real MC background that needs to be subtracted when determining the fake factor. It is found that for the channel $e^{\pm}e^{\pm}$, the hard requirement $M(e_1, e_2) < 50$ GeV makes the region much more fake enriched but also rejects a lot of the data.

Table 23 shows the number of data events, MC events and the average fake factor when applying the two electron measurement regions: $\mu^{\pm}e^{\pm}$ and $e^{\pm}e^{\pm}$ with a hard Z-veto of $M(e_1, e_2) < 50$ GeV. The requirement $N_{\text{b-jets}} = 0$ is applied. The table shows that the measurement region defined by $\mu^{\pm}e^{\pm}$ is much more fake enriched than $e^{\pm}e^{\pm}$ and that the fake factors in the two channels are not mutually consistent. In the measurement regions with $N_{\text{b-jets}} = 0$, fake leptons from the W+jets process are targeted for which the fake lepton should be kinematically independent of the tag lepton and thus not affect the fake factor. The difference in the estimated fake factors could be that the MC modeling of charge misidentification of electrons is not properly scaled or otherwise not sufficient to make an accurate subtraction of the Z-boson events in the $e^{\pm}e^{\pm}$ channel. In order to avoid making a bias for the electron fake factor, only the $\mu^{\pm}e^{\pm}$ channel is used to define the electron measurement region, see Tab. 15.

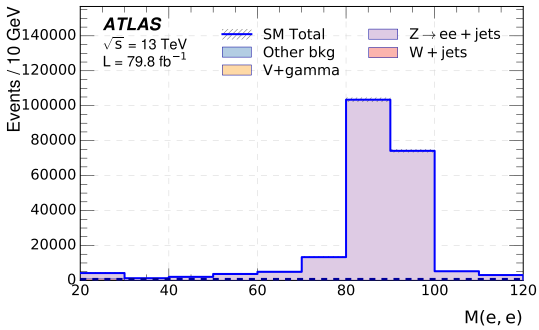
Channel	$p_T(e_{\mathrm{FF}})$	$N_{\mathrm{ID}}^{\mathrm{data}}/10^3$	$N_{\text{anti-ID}}^{\mathrm{data}}/10^3$	$N_{\mathrm{ID}}^{\text{real MC}}/10^3$	$N_{\text{anti-ID}}^{\text{real MC}}/10^3$	F_{average}
$\mu^{\pm}e^{\pm}$	10-50 GeV	42.9±0.2	74.0±0.3	2.1±0.1	0.34±0.04	0.553±0.004
$e^{\pm}e^{\pm}$	10-50 GeV	26.5±0.2	36.9±0.2	5.6±0.3	2.8±0.2	0.602±0.011
$\mu^{\pm}e^{\pm}$	10-20 GeV	35.2±0.2	63.3±0.3	0.84±0.06	0.13±0.02	0.543±0.004
$e^{\pm}e^{\pm}$	10-20 GeV	23.4±0.2	35.3±0.2	3.6±0.3	2.4±0.2	0.60±0.01
$\mu^{\pm}e^{\pm}$	20-30 GeV	4.33±0.07	6.97±0.08	0.51±0.05	0.13±0.03	0.558±0.014
$e^{\pm}e^{\pm}$	20-30 GeV	2.22±0.05	1.29±0.04	1.5±0.2	0.23±0.05	0.64±0.17
$\mu^{\pm}e^{\pm}$	30-40 GeV	2.17±0.05	2.59±0.05	0.46±0.06	0.053±0.02	0.67±0.03
$e^{\pm}e^{\pm}$	30-40 GeV	0.63±0.03	0.22±0.02	0.51±0.05	0.13±0.03	1.4±0.8
$\mu^{\pm}e^{\pm}$	40-50 GeV	1.2±0.05	1.10±0.03	0.33±0.03	0.035±0.002	0.82±0.05
$e^{\pm}e^{\pm}$	40-50 GeV	0.24±0.02	0.054±0.007	0.30±0.02	0.036±0.007	-3.3±2.4

Table 23: Comparison of the number of data events, MC events (with a real e_{FF}) and the average fake factor in the electron measurement regions when requiring a muon as the tag lepton ($\mu^{\pm}e^{\pm}$) and similarly when requiring an electron as the tag lepton ($e^{\pm}e^{\pm}$). The isolation selection `Gradient` is used to define the ID and anti-ID criteria, and $N_{\text{b-jets}} = 0$ is required.

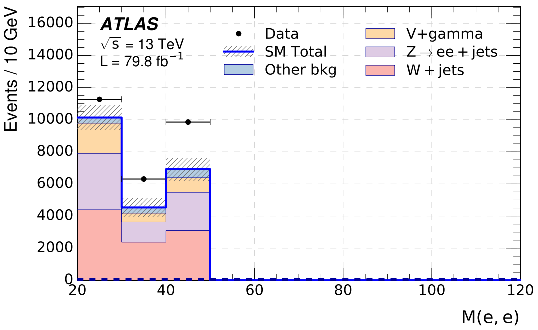
Figure 35 shows the distribution of $p_T(e_{\mathrm{FF}})$ in the electron measurement region. The distribution is showed for both the measurement region where ID and anti-ID criteria are applied and where



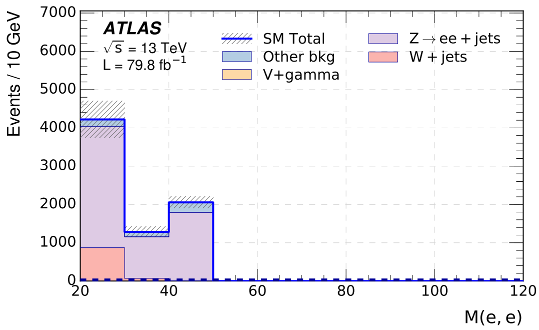
(a) Channel $e^{\pm}e^{\pm}$.
No truth matching requirement.



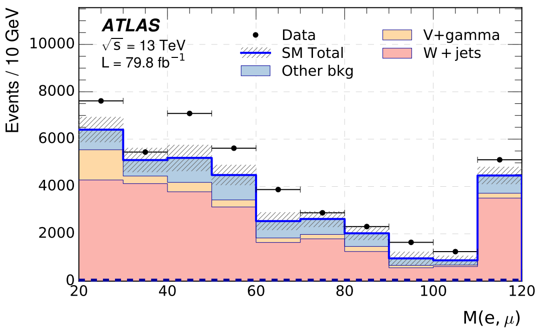
(b) Channel $e^{\pm}e^{\pm}$.
Truth matching requirement: Prompt e_{FF}.



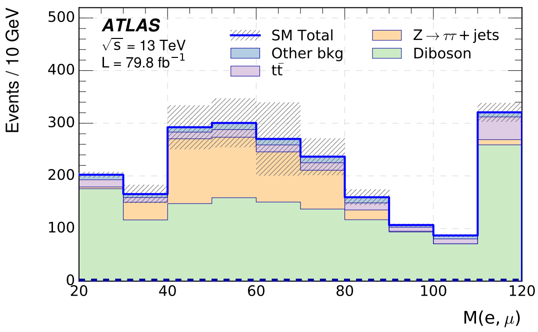
(c) Channel $e^{\pm}e^{\pm}$ with $M(e_1, e_2) < 50$ GeV.
No truth matching requirement.



(d) Channel $e^{\pm}e^{\pm}$ with $M(e_1, e_2) < 50$ GeV.
Truth matching requirement: Prompt e_{FF}.



(e) Channel $\mu^{\pm}e^{\pm}$.
No truth matching requirement.



(f) Channel $\mu^{\pm}e^{\pm}$.
Truth matching requirement: Prompt e_{FF}.

Figure 34: Comparison of the electron measurement region using the channel $e^{\pm}e^{\pm}$ and $\mu^{\pm}e^{\pm}$ with ID criteria for e_{FF} and $N_{\text{b-jets}} = 0$. Distribution of the dilepton invariant mass: $M(e_1, e_2)$ for the $e^{\pm}e^{\pm}$ channel, and $M(\mu, e)$ for the $\mu^{\pm}e^{\pm}$ channel. The blue line indicates the sum of the SM background processes and the dashed area the corresponding statistical uncertainty. The last bin in the histogram includes all events with values larger than the range of the histogram.

truth matching is used to require e_{FF} to be a real electron. It is found that W+jets MC events and $t\bar{t}$ MC events are the primary MC contribution in the electron measurement region for $N_{\text{b-jets}} = 0$ and $N_{\text{b-jets}} \geq 1$ respectively. It is also shown that the region is highly fake enriched with only a small fraction of MC events with a real e_{FF} arising from: diboson events; $t\bar{t}$ events; and $Z \to \tau^{\pm}\tau^{\mp}$ events. Finally, it is shown that the electron measurement region with ID and anti-ID criteria targets similar MC events.

In the signal region, very few MC events have a fake electron. The low statistics for fake electrons in the signal region results in very large uncertainties on the composition study which showed that the composition is a mix of fake electrons from heavy flavor decays labeled "non-isolated", electron fakes which originate from light flavor jets labeled "hadrons", and electron fakes from photon conversion and decays of light mesons labeled "bkg lep".

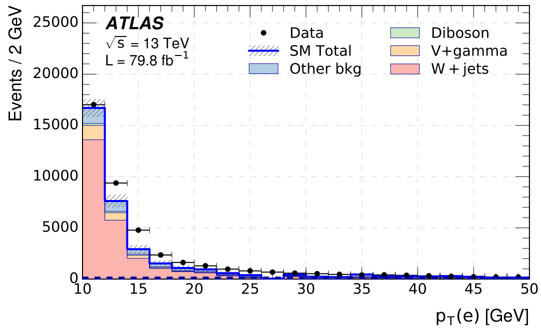
The composition of e_{FF} in the electron MR is studied in Fig. 36. The truth matching for e_{FF} in the MC events shows that the composition of fake e_{FF} is also a mix of electrons from the three fake sources: "non-isolated"; "hadrons", and "bkg lep". When e_{FF} satisfy the anti-ID criteria, the fraction of events arising from electron fakes originating from light flavor jets ("hadrons") increases slightly. Ideally, the composition of the fake electron events should be similar in the numerator and denominator of the fake factor. Since the fake electrons of truth type "hadrons" tend to be less isolated than the fake electrons of truth type "non-isolated", the difference in the fake composition can result in slightly smaller fake factors. In Sec. 9 are systematic uncertainties of the model examined where differences in the fake composition are taken into account.

Due to the limited statistics for the composition of fake electron events in the signal region, it is not possible to determine how well the composition of fake electrons in the signal region matches the composition in the measurement region. The kinematic differences between the three types of fake composition ("non-isolated", "hadrons" and "bkg lep") are studied using the electron measurement region. The study is included in appendix Sec. E. It is found that the fake electrons of type "non-isolated" tend to be the most isolated and the fake electrons of type "hadrons" the least isolated. Furthermore, the fake electrons of type "non-isolated" tend to have a higher p_T, $|d_0|/\sigma(d_0)$ and $|\Delta z_0 \sin(\theta)|$. This shows that the necessary binning of the fake factor will depend on the electron composition.

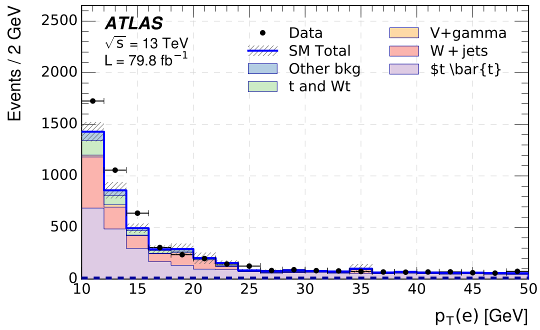
Table 24 provides an overview of the number of data events, number of MC events with a real e_{FF} and the average fake factor in each of the electron MRs. It is shown that the `Loose` isolation is worse at discriminating fake and real leptons than the `Gradient` isolation as expected.

	$N_{\mathrm{ID}}^{\mathrm{data}}/10^3$	$N_{\text{anti-ID}}^{\mathrm{data}}/10^3$	$N_{\mathrm{ID}}^{\text{real MC}}/10^3$	$N_{\text{anti-ID}}^{\text{real MC}}/10^3$	F_{average}
`FixedCutLoose`	48.2±0.2	84.9±0.3	2.9±0.1	0.51±0.04	0.54
`Loose`	78.0±0.3	55.0±0.2	3.3±0.1	0.14±0.02	1.36

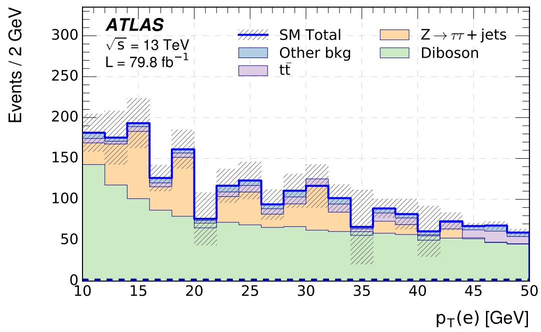
Table 24: Number of data events, number of MC events (with a real μ_{FF}) and the average fake factor in the applied electron measurement regions. The isolation selections used to determine the ID and anti-ID criteria are listed.



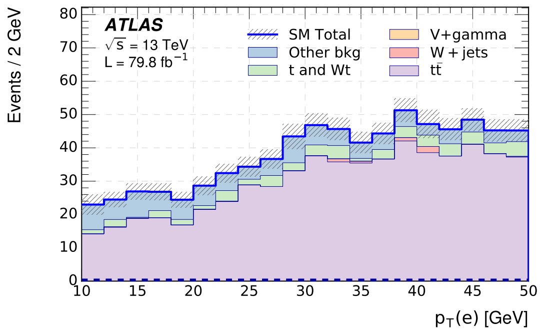
(a) Electron MR. ID criteria. $N_{\text{b-jets}}$=0.
No truth matching requirement.



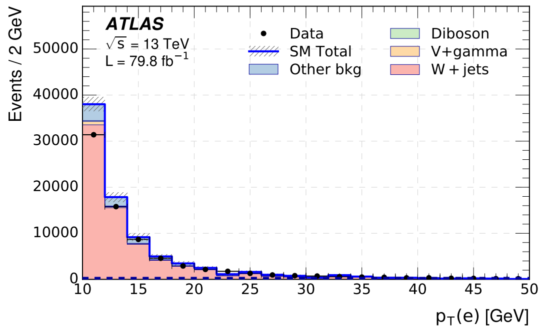
(b) Electron MR. ID criteria. $N_{\text{b-jets}} \geq 1$.
No truth matching requirement.



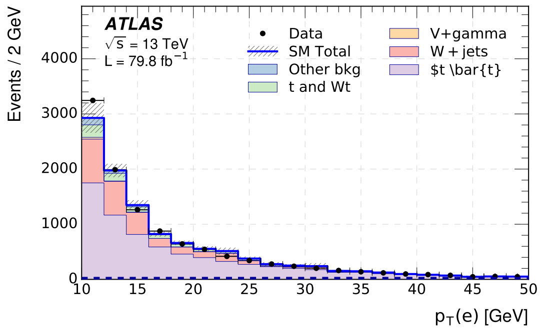
(c) Electron MR. ID criteria. $N_{\text{b-jets}}$=0.
Truth matching requirement: Prompt e_{FF}.



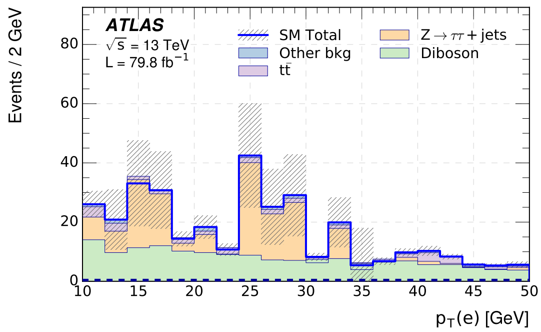
(d) Electron MR. ID criteria. $N_{\text{b-jets}} \geq 1$.
Truth matching requirement: Prompt e_{FF}.



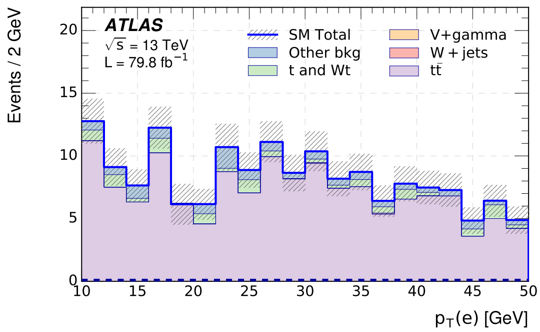
(e) Electron MR. Anti-ID criteria. $N_{\text{b-jets}}$=0.
No truth matching requirement.



(f) Electron MR. Anti-ID criteria. $N_{\text{b-jets}} \geq 1$.
No truth matching requirement.

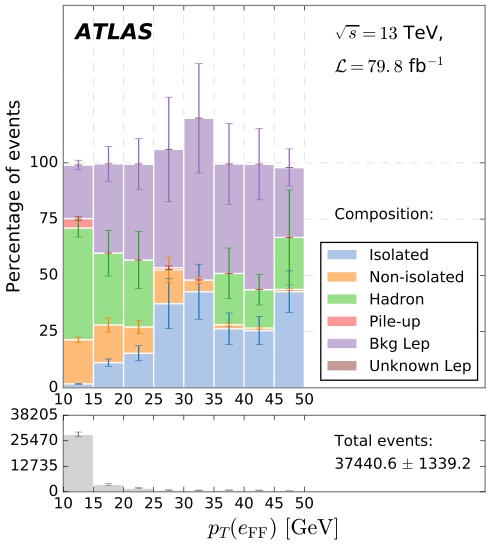


(g) Electron MR. Anti-ID criteria. $N_{\text{b-jets}}$=0.
Truth matching requirement: Prompt e_{FF}.

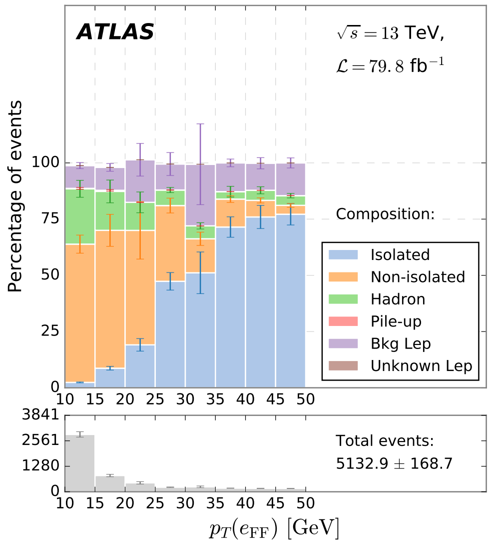


(h) Electron MR. Anti-ID criteria. $N_{\text{b-jets}} \geq 1$.
Truth matching requirement: Prompt e_{FF}.

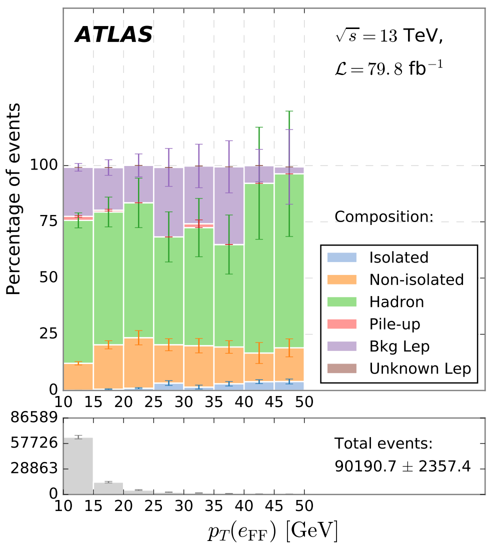
Figure 35: The electron measurement region with ID and anti-ID criteria for e_{FF}. Distribution of $p_T(e_{\text{FF}})$. The blue line indicates the sum of the SM background processes and the dashed area the corresponding statistical uncertainty. The last bin in the histogram includes all events with values larger than the range of the histogram.



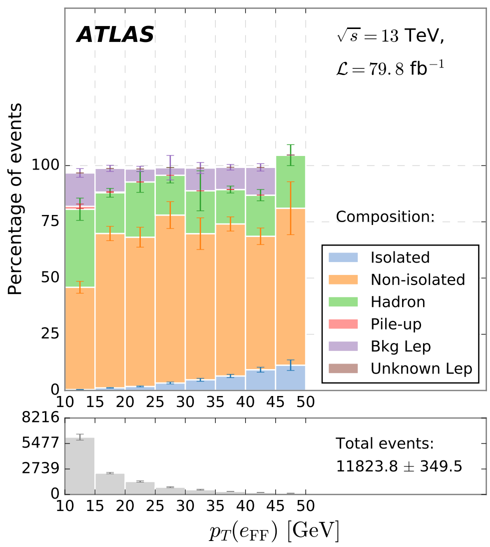
(a) ID criteria for e_{FF}. $N_{\text{b-jets}}$=0.



(b) ID criteria for e_{FF}. $N_{\text{b-jets}} \geq 1$.



(c) Anti-ID criteria for e_{FF}. $N_{\text{b-jets}}$=0.



(d) Anti-ID criteria for e_{FF}. $N_{\text{b-jets}} \geq 1$.

Figure 36: Composition of e_{FF} in the electron MR as a function of $p_T(e_{\mathrm{FF}})$ when requiring e_{FF} to satisfy the ID criteria (top plots) and the anti-ID criteria (bottom plots). The final bin contains all the events with $p_T > 45$ GeV. The main plots show the composition as the percentage of events in the p_T bin, whereas the subplots show the number of events in each p_T bin and the total number of events.

8.3.8 Binning of the Electron Fake Factor

In this section, the binning applied for the electron fake factor is determined. The isolation selection `Gradient` is used for e_{FF}. The procedure used in this section is similar to the one applied in Sec. 8.3.6 for the muon fake factors.

Figure 37 shows how the fake factor depend on $N_{\text{b-jets}}$. Again, it is found that the fake factor has a clear dependency on $N_{\text{b-jets}}$, so the fake factor is divided into the two regions: $N_{\text{b-jets}} = 0$ and $N_{\text{b-jets}} \geq 1$. The fake factor determined in the $N_{\text{b-jets}} = 0$ measurement region will be used to determine the data-driven fake background in the signal region. The fake factors determined in the $N_{\text{b-jets}} \geq 1$ measurement region are constructed too in order to check the stability of the model in a control region with $N_{\text{b-jets}} \geq 1$. In the following section, I will focus on the fake factor for which $N_{\text{b-jets}} = 0$.

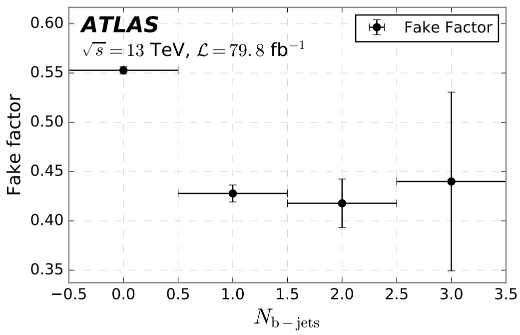
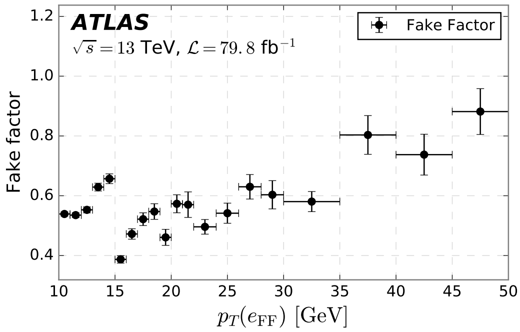


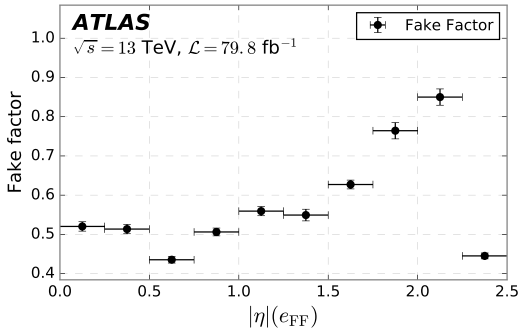
Figure 37: Electron fake factor as a function of $N_{\text{b-jets}}$. The ID and anti-ID criteria are defined using the `Gradient` isolation.

Figure 38 shows the fake factor as a function of these six variables: $p_T(e_{\mathrm{FF}})$, $|\eta|(e_{\mathrm{FF}})$, $|d_0|/\sigma(d_0)(e_{\mathrm{FF}})$, $|\Delta z_0 \sin(\theta)|(e_{\mathrm{FF}})$, N_{jets}, and E_T^{miss}. The following figures used in this study are included in appendix Sec. F. Figure 121 shows in detail the fake factor dependency on the transverse momentum of the fake factor electron. The shape of the fake factor as a function of $p_T(e_{\mathrm{FF}})$ reflects that the ID and anti-ID criteria are defined using an efficiency based isolation selection based on a cut map. This can make it more difficult to determine the appropriate binning since the fake factor distribution can make rapid jumps when the efficiency requirements change. The majority of the fake electrons have small $p_T(e_{\mathrm{FF}})$, which makes it possible to apply relatively small bin widths in this area and thereby get a more accurate description of the fake factor dependency on $p_T(e_{\mathrm{FF}})$ without having too low statistical uncertainties in each bin. The bins are defined in the following ranges of $p_T(e_{\mathrm{FF}})$: 10-13, 13-15, 15-17, 17-35, and 35-50 GeV.

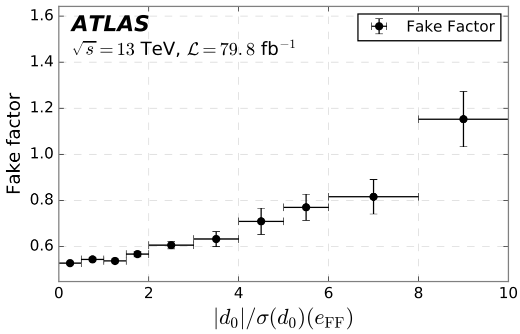
The fake factor dependency of the variables $|d_0|/\sigma(d_0)$, $|\Delta z_0 \sin(\theta)|$ and $|\eta|$ for e_{FF} and E_T^{miss} and N_{jets} are examined again for the fake factor in each of the five $p_T(e_{\mathrm{FF}})$ bins. Figure 122 shows the fake factor as a function of $|d_0|/\sigma(d_0)$ and $|\Delta z_0 \sin(\theta)|$ for e_{FF}. For small $p_T(e_{\mathrm{FF}})$, the fake factor increase when the impact parameter variables increase. Figure 123 shows the fake factor as a function of $|\eta|(e_{\mathrm{FF}})$ and E_T^{miss}. In all the $p_T(e_{\mathrm{FF}})$ bins, the fake factor tend to increase when $|\eta| > 1.5$. The fake factor dependence on E_T^{miss} is not showing a clear tendency across the five bins in $p_T(e_{\mathrm{FF}})$. Neither $|\eta|(e_{\mathrm{FF}})$ or E_T^{miss} are used to bin the fake factor in lack of a clear shape dependency. Figure 124 shows the fake factor as a function of N_{jets} for each of the five $p_T(e_{\mathrm{FF}})$ bins. It is found that for all the p_T bins, the fake factor decreases when the number of jets increase. In order to get the best description of the fake factor dependencies, it is chosen to apply bins as a function of N_{jets} and afterwards consider the fake factor distribution as a function of $|d_0|/\sigma(d_0)(e_{\mathrm{FF}})$ to ensure a good description of the fake factor shape for low p_T fake electrons. Two bins are applied as a function of N_{jets}: $N_{\text{jets}} \leq 2$ and $N_{\text{jets}} > 2$.



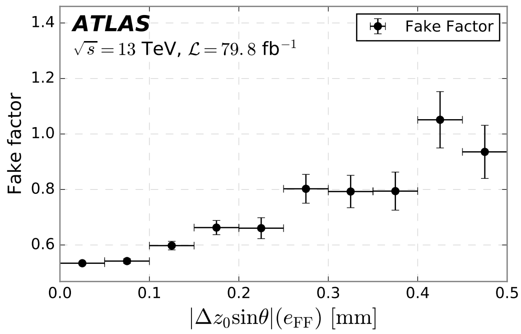
(a) Fake factor as a function of $p_T(e_{\mathrm{FF}})$.



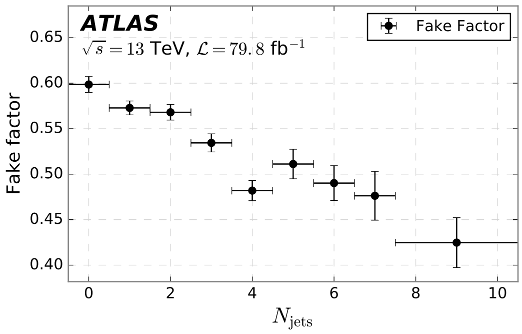
(b) Fake factor as a function of $|\eta|(e_{\mathrm{FF}})$.



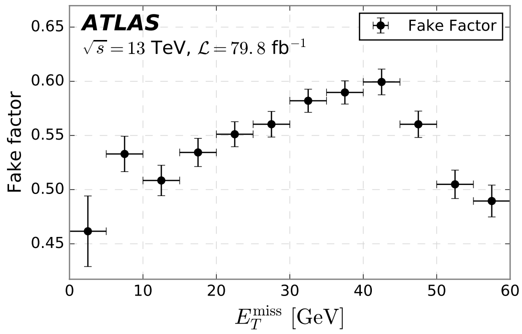
(c) Fake factor as a function of $|d_0|/\sigma(d_0)(e_{\mathrm{FF}})$.



(d) Fake factor as a function of $|\Delta z_0 \sin(\theta)|(e_{\mathrm{FF}})$.



(e) Fake factor as a function of N_{jets}.

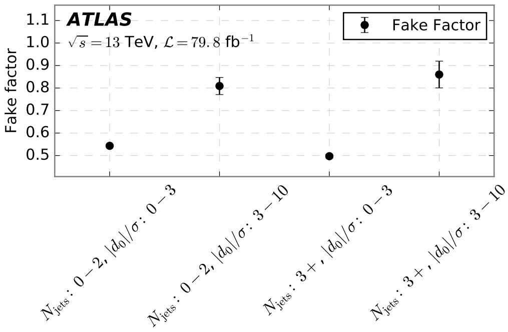


(f) Fake factor as a function of E_T^{miss}.

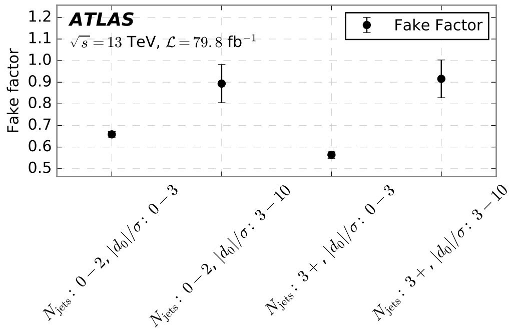
Figure 38: Electron fake factor as a function of: $p_T(e_{\mathrm{FF}})$; $|\eta|(e_{\mathrm{FF}})$; $|d_0|/\sigma(d_0)(e_{\mathrm{FF}})$; $|\Delta z_0 \sin(\theta)|(e_{\mathrm{FF}})$; N_{jets}; and E_T^{miss}. The ID and anti-ID criteria are defined using the `Gradient` isolation in the electron measurement region with $N_{\text{b-jets}} = 0$.

The fake factor dependency on $|d_0|/\sigma(d_0)(e_{\mathrm{FF}})$ is examined in Fig. 125 for the fake factor bin $N_{\mathrm{jets}} \leq 2$ and the five bins in $p_T(e_{\mathrm{FF}})$ and similarly in Fig. 126 for $N_{\mathrm{jets}} > 2$. Two bins are applied: $|d_0|/\sigma(d_0)(e_{\mathrm{FF}}) < 3$ and $3 \leq |d_0|/\sigma(d_0)(e_{\mathrm{FF}}) < 10$. The final bins applied for the electron fake factor are shown in Fig. 39. It is found that the two bins in $|d_0|/\sigma(d_0)(e_{\mathrm{FF}})$ are consistent within the statistical uncertainties for the three fake factor bins with largest $p_T(e_{\mathrm{FF}})$ and therefore merged, so bins in $|d_0|/\sigma(d_0)(e_{\mathrm{FF}})$ are only applied for the fake factor with lowest $p_T(e_{\mathrm{FF}})$ as expected.

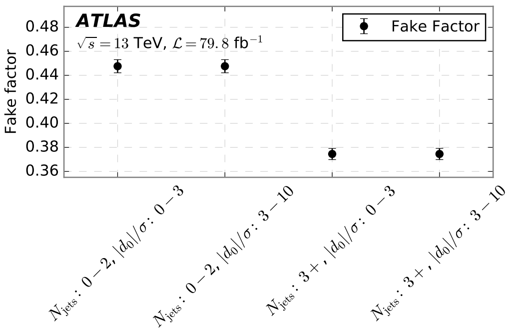
A study of the fake factor dependencies is similarly made for the electron measurement region for which $N_{\text{b-jets}} \geq 1$ and the measurement regions applying looser isolation selection. Studied showing the fake factor dependencies in these measurement regions are included in appendix Sec. F. Table 25 shows the applied binning of the electron fake factor in all the considered measurement regions



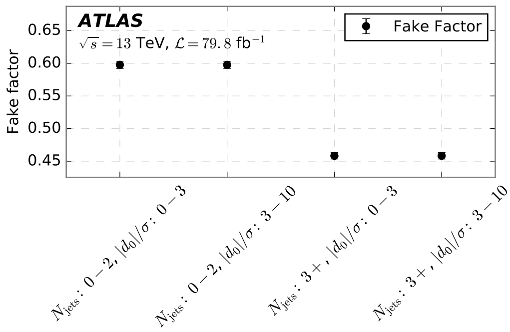
(a) Bin: 10 GeV $\leq p_T(e_{\text{FF}}) < 13$ GeV



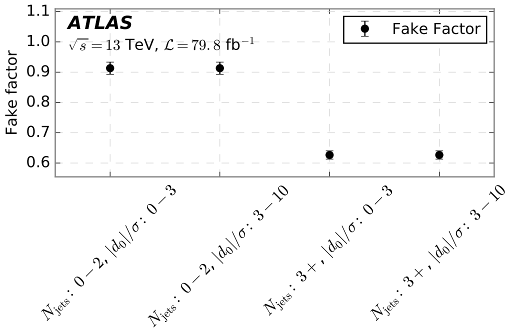
(b) Bin: 13 GeV $\leq p_T(e_{\text{FF}}) < 15$ GeV



(c) Bin: 15 GeV $\leq p_T(e_{\text{FF}}) < 17$ GeV



(d) Bin: 17 GeV $\leq p_T(e_{\text{FF}}) < 35$ GeV



(e) Bin: 35 GeV $\leq p_T(e_{\text{FF}}) < 50$ GeV

Figure 39: The final binning for the electron fake factor. The ID and anti-ID criteria are defined using the `Gradient` isolation in the electron measurement region with $N_{\text{b-jets}} = 0$.

defined by the isolation selection for e_{FF} and the $N_{\text{b-jets}}$ requirement. All additional requirements for the measurement regions are constant, as listed in Tab. 15.

8.4 Control Regions

The stability of the fake prediction against data is studied in three control regions which are kinematically similar to the signal region. It is necessary to use control regions to check the agreement between data and the estimated background since data can not be applied in the "blinded" signal region in order to ensure that a bias is not introduced in the background method. The control regions should therefore be similar but independent to the signal region so no signal is expected in these regions while the kinematics are similar. The control regions are also used to estimate the systematic uncertainties on the fake prediction in the following section.

Measurement region		Applied fake factor bins		
Isolation	$N_{\text{b-jets}}$	$p_T(\mu_{\text{FF}})$	N_{jets}	$\lvert d_0\rvert/\sigma(d_0)(\mu_{\text{FF}})$
`FixedCutLoose`	0	5 bins	2 bins	2 bins
`FixedCutLoose`	≥ 1	3 bins	–	3 bins
`Loose`	0	5 bins	2 bins	2 bins
`Loose`	≥ 1	3 bins	2 bins	2 bins

Table 25: The applied fake factor bins in each of the electron measurement regions. The measurement regions are defined by the isolation selection used for e_{FF} which defines the ID and anti-ID criteria, see Tab. 20, and the requirement for $N_{\text{b-jets}}$. All the final fake factor bins are included in appendix Sec. F.

Three control regions are defined as neighboring the signal region by inverting exactly one signal requirement and loosening the charge requirement of the electron to also include fake lepton events from the all same-sign region. By using several control regions, the available statistical size of background samples is increased and it is better examined if the background estimation is stable. Table 26 shows the requirements applied to the three control regions.

Control region	Lepton charge requirement	Inverted signal region requirement
1	$\mu^\pm\mu^\pm e^\mp$ + $\mu^\pm\mu^\pm e^\pm$	$M(\mu_1,\mu_2,e) < 40$ GeV or $M(\mu_1,\mu_2,e) > 90$ GeV
2	$\mu^\pm\mu^\pm e^\mp$ + $\mu^\pm\mu^\pm e^\pm$	$N_{\text{b-jets}} \geq 1$
3	$\mu^\pm\mu^\pm e^\mp$ + $\mu^\pm\mu^\pm e^\pm$	$E_T^{\text{miss}} > 60$ GeV

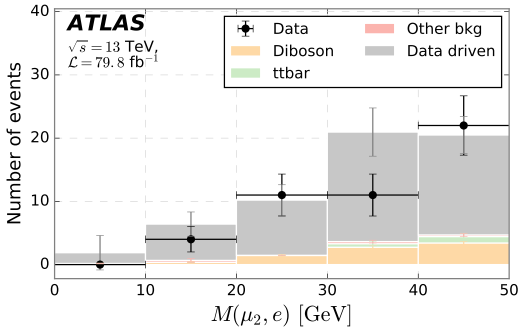
Table 26: Definitions of the three control regions. Excpept for the inverted requirement, all the signal region requirements are applied in the control regions, see Tab. 11

Control region 1 is defined by inverting the trilepton invariant mass requirement. The trilepton mass requirement is applied in the signal region due to the correlation between the W-boson mass and the trilepton final state in the signal samples and not targeting any specific background processes. This requirement is thus not changing the type of background processes significantly from the background processes seen in the signal region. *Control region 2* is a b-jet enriched region. The region has a big contribution from the $t\bar{t}$ background process. In the previous sections, it was found that the presence of b-jets in the measurement region changes the shape and normalization of the fake factors. Separate fake factors were measured in b-jet enriched measurement regions which will be applied to estimate the fake background in control region 2. *Control region 3* is defined by inverting the E_T^{miss} requirement. The background processes which have large E_T^{miss} have either one or more energetic SM neutrinos in the final state that arise from Z- or W-boson decays. The control region is thus enriched in $t\bar{t}$ and diboson background. Pile-up or miss-reconstructed final state particles can also contribute to the E_T^{miss}. In the previous sections, it was found that the fake factors did not have a large dependence on E_T^{miss} when requiring $E_T^{\text{miss}} < 60$ GeV. This does not necessarily mean that the fake factors do not have a large E_T^{miss} dependence for $E_T^{\text{miss}} \geq 60$ GeV. In this analysis, a possible E_T^{miss} dependency for the fake factor has not been taken into account for the data-driven model. The data-driven estimate is thus expected to be a less accurate reflection of the fake lepton kinematics in control region 3 compared to control region 1 and 2.

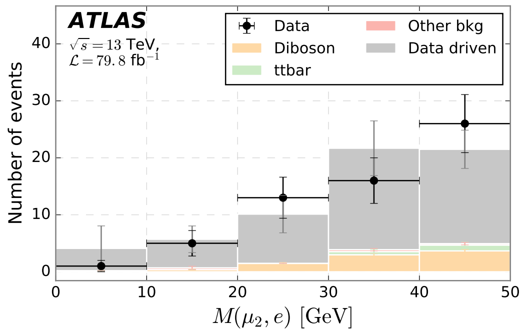
A composition study is made for the final state leptons in the three control regions which is shown in appendix Sec. H. The dominant electron fake type varies for the three control regions as expected. In control region 1, fakes of type "hadrons" are the dominant contribution, whereas in control region 2 and 3, fakes of type "non-isolated" are dominating. For all the control regions is the fake electron type originating from a combination of "non-isolated", "hadron" and "bkg lep", which is also the case in the signal region and electron measurement region. The muon fake type is almost exclusively of type "non-isolated". A small contribution from "pile-up" muons is also present in all the control regions.

Figure 40-42 show the distribution of $M(\mu_2, e)$ in control region 1, 2 and 3 respectively. The background estimation is consistent with data when applying the `FixedCutLoose` or `Loosened FixedCutLoose` isolation for the subleading muon and the `Gradient` or `Loose` isolation for the electron, where the background estimation performs badly when the `Loose` isolation selection is applied for the subleading muon. The `Loose` isolation selection is very bad at discriminating real and fake muons, so a lot of fake muons satisfy this requirement. This results in very large fake factors. In the previous section is was found that the average muon fake factor when applying the `Loose` isolation selection is 5.8 where it is 0.35 and 0.57 respectively when the `FixedCutLoose` and `Loosened FixedCutLoose` isolation selections are used. The data-driven in regions using the `Loose` isolation selection for muons is thus based on fewer events satisfying the anti-ID criteria which are multiplied by a very high fake factor, see Eq. 8.8. Due to the very bad background modeling when applying the `Loose` isolation selection for the subleading muon, this region will not be applied in the next section to estimate systematic uncertainties.

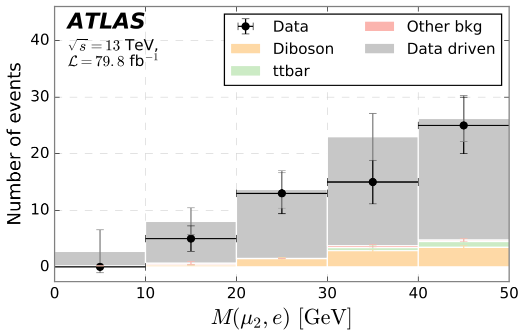
There is a larger discrepancy between data and the background estimate in control region 3 compared to control region 1 and 2. This could be caused by the fake factors not being representative for the background processes in this region. The fake factor dependency on E_T^{miss} in the region with $E_T^{\text{miss}} > 60$ GeV should therefore be studied in future work when applying this control region.



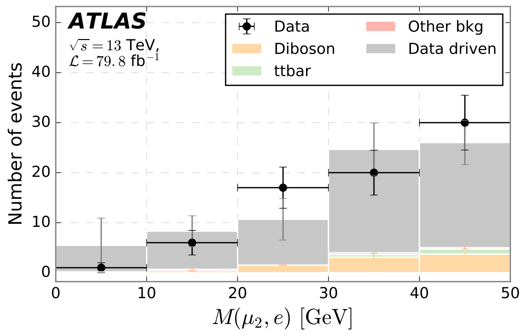
(a) μ_2: `FixedCutLoose` isolation.
e: `Gradient` isolation.



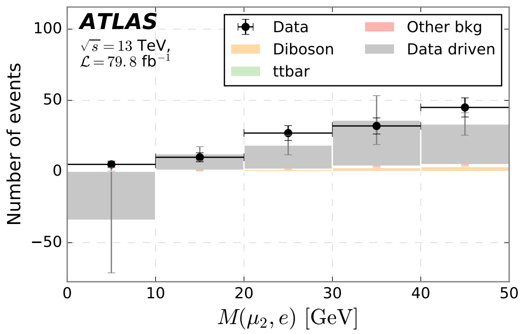
(b) μ_2: `FixedCutLoose` isolation.
e: `Loose` isolation.



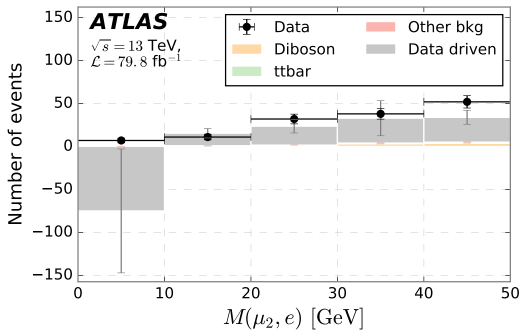
(c) μ_2: `Loosened FixedCutLoose` isolation.
e: `Gradient` isolation.



(d) μ_2: `Loosened FixedCutLoose` isolation.
e: `Loose` isolation.

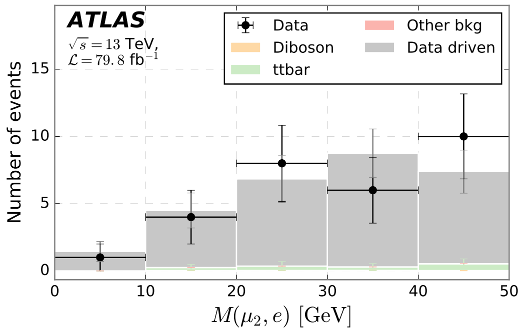


(e) μ_2: `Loose` isolation.
e: `Gradient` isolation.

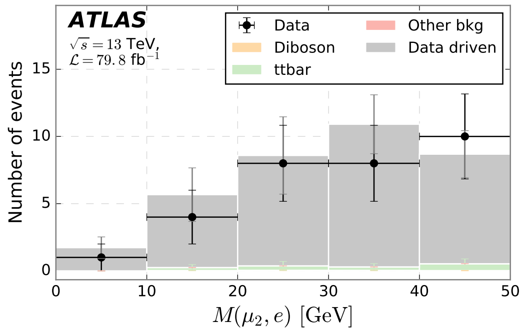


(f) μ_2: `Loose` isolation.
e: `Loose` isolation.

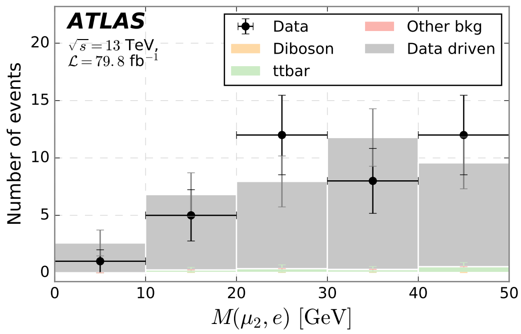
Figure 40: Control Region 1 - Inverted $M(\mu_1, \mu_2, e)$ requirement. Distribution showing $M(\mu_2, e)$.



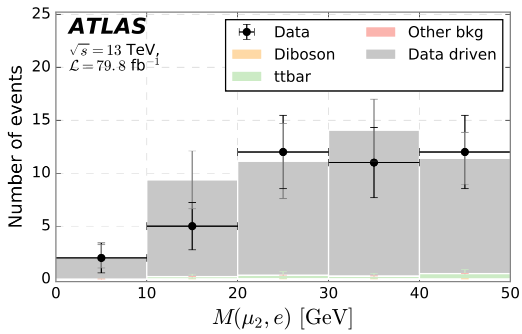
(a) μ_2: `FixedCutLoose` isolation.
e: `Gradient` isolation.



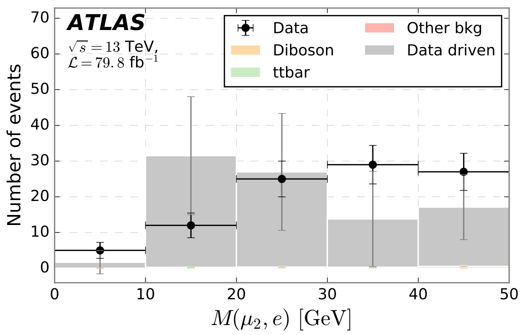
(b) μ_2: `FixedCutLoose` isolation.
e: `Loose` isolation.



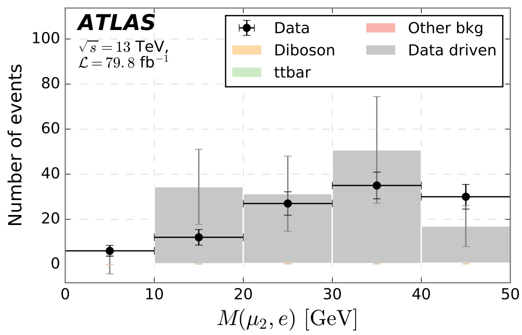
(c) μ_2: `Loosened FixedCutLoose` isolation.
e: `Gradient` isolation.



(d) μ_2: `Loosened FixedCutLoose` isolation.
e: `Loose` isolation.

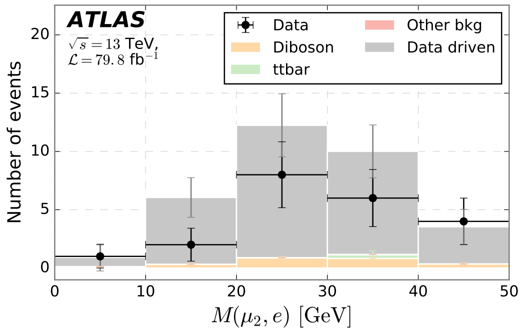


(e) μ_2: `Loose` isolation.
e: `Gradient` isolation.

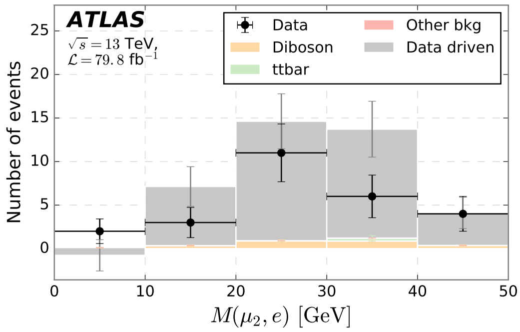


(f) μ_2: `Loose` isolation.
e: `Loose` isolation.

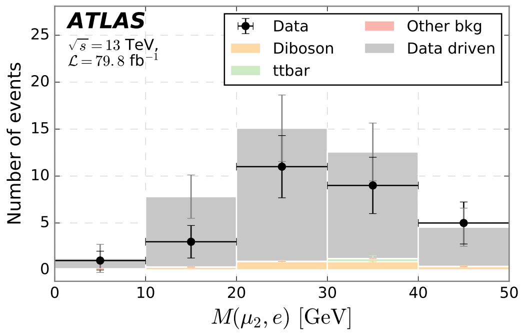
Figure 41: Control Region 2 - $N_{\text{b-jets}} \geq 1$ requirement. Distribution showing $M(\mu_2, e)$.



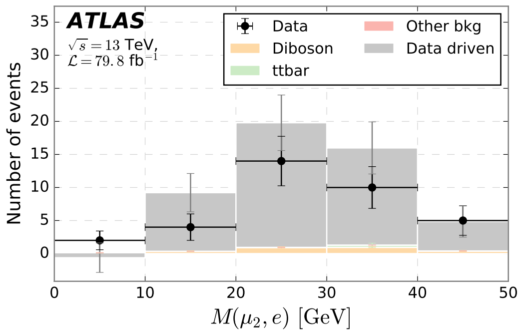
(a) μ_2: `FixedCutLoose` isolation.
e: `Gradient` isolation.



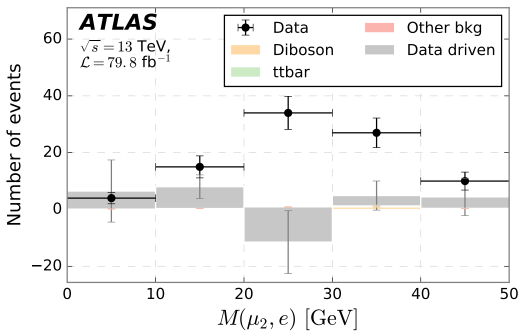
(b) μ_2: `FixedCutLoose` isolation.
e: `Loose` isolation.



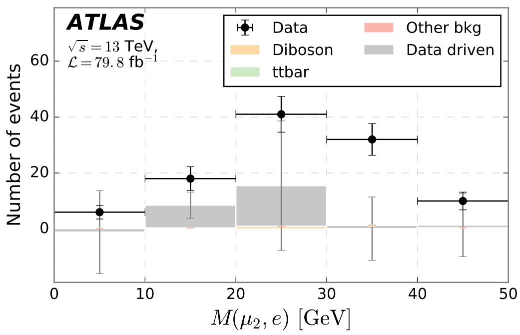
(c) μ_2: `Loosened FixedCutLoose` isolation.
e: `Gradient` isolation.



(d) μ_2: `Loosened FixedCutLoose` isolation.
e: `Loose` isolation.



(e) μ_2: `Loose` isolation.
e: `Gradient` isolation.



(f) μ_2: `Loose` isolation.
e: `Loose` isolation.

Figure 42: Control Region 3 - Inverted E_T^{miss} requirement. Distribution showing $M(\mu_2, e)$.

9 Systematic Uncertainties

Several sources of systematic uncertainties affect the background prediction. Experimental uncertainties arise from the modeling of particle reconstruction in the detector simulation, the luminosity uncertainty and pile-up measurement, and the systematic effects for the data-driven model. Theoretical uncertainties arise from: the simulations of the SM background processes and signal process including the calculation of the cross-sections, the applied parton distribution functions, and the renormalization and factorization scales governing the QCD coupling strength and the separation of the hard and soft scale contributions in the cross section [1]. At the time of writing, the ATLAS analysis group [15], which is doing the 2016 search for prompt HNLs, is in contact with theorists regarding the calculation of the HNL cross section which can change the normalization of the signal samples applied. This can also change the theoretical uncertainties for the analysis.

In this analysis, statistical uncertainties are applied for both the data and MC samples. Systematic uncertainties are estimated for the data-driven method in this section. These systematic uncertainties will be applied to the fake factors when estimating the data-driven background. Additional systematic uncertainties for the MC samples are not applied. Since the large majority of the SM background processes are estimated using the data-driven model, it is expected that the systematic uncertainties related to the MC contribution will have a much smaller impact on the overall background estimation. In a full analysis of the HNL signal, all the systematic uncertainties mentioned above should be applied in the background estimation.

9.1 Fake factor uncertainties

The systematic uncertainties of the fake factor model are examined in this section. The assigned systematic uncertainties will be applied to the muon and electron fake factors and will thereby affect the total uncertainty on the data-driven estimate through the error propagation. In order to get the total systematic uncertainty, the individual systematic uncertainties are added in quadrature.

The following sources of systematic uncertainties are addressed.

Model stability	:	The stability of the background estimation is examined by comparing the agreement of data and the background estimate using different ID and anti-ID criteria in the three control regions. The varying fake lepton composition in the three control regions allows to assign a systematic uncertainty to the fact that the fake composition in the signal region and measurement region is not exactly the same.
Subtracted MC	:	Prompt MC is subtracted from data when determining the fake factor. The technique therefore relies on the MC being an accurate model of data. An uncertainty is assigned based on how much the fake factors change when relevant MC processes are scaled up/down.

Model stability

The agreement between data and the background estimation is examined in this section. Two separate things are varied in this study: the control region applied which changes the fake lepton composition; and the ID and anti-ID criteria used, i.e. changing the isolation selection, which affects the statistics both directly in the region and for the data-driven model. The looser ID and anti-ID criteria were described in Sec. 8.3.4.

The three control regions introduced in Sec. 8.4 are used to determine the agreement between data and the background estimation. In the previous section, it was found that the fake lepton composition changes for the three control regions. In control region 1, the fake electrons are mainly fakes from light flavor hadrons, whereas the fake electrons mainly arise from semi-leptonic heavy flavor decays in control region 2 and 3. The stability of the background estimation with respect to changes in the background composition is thus tested when comparing data to the predicted data-driven estimate in the three regions.

The three “loose selections” used in the study are listed in Tab. 27. They are used to define the ID and anti-ID criteria applied in the fake factor model. A looser isolation selection will allow more events in the control region but less events in the anti-ID region used to determine the fake background estimate, see Eq. 8.8.

	Isolation selections used to define the ID and anti-ID criteria			
	Signal selection	Loose selection 1	Loose selection 2	Loose selection 3
Isolation for μ_2	`FixedCutLoose`	`FixedCutLoose`	`Loosened FixedCutLoose`	`Loosened FixedCutLoose`
Isolation for e	`Gradient`	`Loose`	`Gradient`	`Loose`

Table 27: The isolation selections used in the signal region requirements and the three regions used to check the stability of the background estimation referred to as “loose region” 1, 2 and 3.

The data and background agreement is tested in the two distributions: $p_T(\mu_2)$ and $p_T(e)$, which are sensitive to the applied isolation selections. Since the data change when changing the control region or isolation criteria, the ratio of the number of data events and estimated background events is considered:

$$R = \frac{N_{\text{data}}}{N_{\text{Bkg}}}. \tag{9.1}$$

Figure 43 shows the ratio R as a function of $p_T(\mu_2)$ in control region 1 for the four different isolation selections listed in Tab. 27. Similar figures are included in appendix Sec. I.1 which shows R as a function of $p_T(\mu_2)$ and $p_{\text{T}}(e)$ for the three control regions.

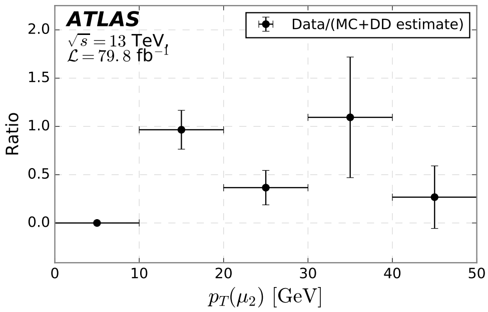
The model stability is tested using the following variable denoted S which quantifies the average squared difference of R:

$$S = \frac{1}{N_{\text{bins}}} \left[\sum_{i}^{N_{\text{bins}}} \left(R^{\text{control region}}_{\text{signal selection}}(i) - R^{\text{control region}}_{\text{loose selection}}(i) \right)^2 \right]^{1/2}, \tag{9.2}$$

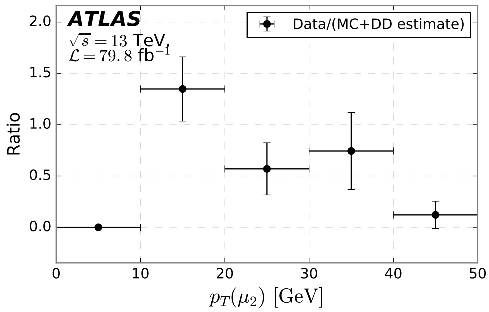
where $N_{\text{bins}} = 4$ for the considered distributions, $p_T(\mu_2)$ and $p_{\text{T}}(e)$. Table 28 shows S for the three control regions for each of the isolation selections listed in Tab. 27. The obtained result for control region 3 using the $p_T(\mu_2)$ distribution is caused by very small statistics in the 30 GeV $< p_T(\mu_{\text{FF}}) \leq 40$ GeV bin, see Fig. 148 in appendix Sec. I.1, and is therefore not applied when assigning an uncertainty.

The variable S is affected by both changing statistics, due to the varied isolation criteria and control regions, and by the changing fake lepton composition in the control regions. One combined modeling uncertainty is thus determined based on the study. The maximum value of S is 0.21, so a systematic uncertainty of 21% is used to the fake factors as a conservative estimate of the uncertainties related to the model stability against changing statistics through the definition of ID and anti-ID criteria and variations in the fake composition between the signal region and measurement region.

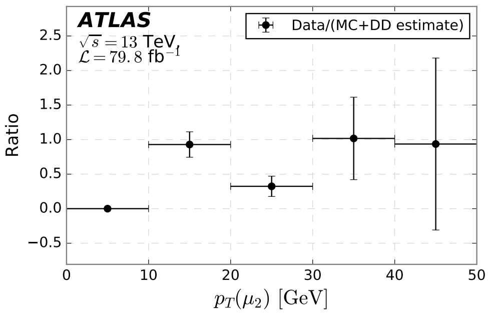
The assigned systematic uncertainty is constant for the fake factors in all bins. In future work, it should be considered whether the systematic uncertainty should vary as a function of e.g. p_T of the fake lepton.



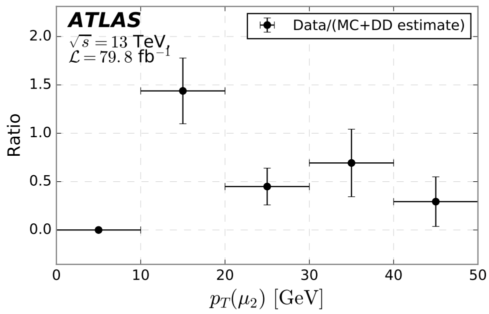
(a) μ_2: `FixedCutLoose` isolation.
e: `Gradient` isolation.



(b) μ_2: `FixedCutLoose` isolation.
e: `Loose` isolation.



(c) μ_2: `Loosened FixedCutLoose` isolation.
e: `Gradient` isolation.



(d) μ_2: `Loosened FixedCutLoose` isolation.
e: `Loose` isolation.

Figure 43: Control Region 1 - Inverted $M(\mu_1, \mu_2, e)$ requirement. Distribution showing $p_T(\mu_2)$.

Control region	Distribution	Loose selection 1	Loose selection 2	Loose selection 3
1	$p_T(\mu_2)$	0.14±0.61	0.17±1.7	0.16±0.69
1	$p_T(e)$	0.12±0.49	0.12±0.57	0.21±0.98
2	$p_T(\mu_2)$	0.05±0.23	0.08±0.49	0.08±0.43
2	$p_T(e)$	0.12±0.57	0.13±0.75	0.20±0.98
3	$p_T(\mu_2)$	2±570	3±857	2±415
3	$p_T(e)$	0.02±0.12	0.16±1.6	0.14±1.4

Table 28: Comparison of the modeling stability defined by S, see Eq. 9.2. The three loose selections are defined in Tab. 27. For both the $p_T(\mu_2)$ and $p_{\mathrm{T}}(e)$ distribution $N_{\mathrm{bins}} = 4$ in the p_T range from 10-50 GeV.

Subtracted MC

The systematic uncertainties introduced by the subtraction of MC samples is now examined. In analyses similar to this one, it is found that a systematic uncertainty can be introduced by the subtraction of diboson MC samples where the probe lepton is real [62, 63]. In order to determine how big an influence this can have, a rescale factor is used for the MC subtraction from diboson samples which is varied up and down by 15%. Similar analyses as [62, 63] apply a 15% scaling of the subtracted MC as a conservative estimate of how much the normalization factors. The resulting fake factors are compared to the nominal fake factors by determining the percentage difference:

$$P_{e,VV\pm15\%} = \frac{F(e_{\mathrm{FF}}) - F_{VV\pm15\%}(e_{\mathrm{FF}})}{F(e_{\mathrm{FF}})} \cdot 100\%,$$
$$P_{\mu,VV\pm15\%} = \frac{F(\mu_{\mathrm{FF}}) - F_{VV\pm15\%}(\mu_{\mathrm{FF}})}{F(\mu_{\mathrm{FF}})} \cdot 100\%. \tag{9.3}$$

Table 77 included in appendix Sec. I.2 shows the percentage difference between the fake factors with and without the 15% scaling of the diboson samples. It is found that the largest difference is 10% and 5% for the muon and electron fake factors respectively. The systematic uncertainties governing the MC subtraction are therefore chosen as 10% and 5% for the muon and electron fake factors respectively.
The two systematic uncertainties for the fake factor methods are added in quadrature to determine the total fake factor uncertainty. The total systematic uncertainty for the muon and electron fake factors is 23% and 22% respectively. In order to get the total uncertainty for each fake factor bin is the assigned systematic uncertainty added in quadrature with the statistical uncertainty. When determining the signal region requirements based on MC modeling of the background processes in Sec. 7.3, a systematic uncertainty of 30% and 50% was applied to the background modeling. This estimate was thereby roughly comparable to the systematic uncertainty assigned to the data-driven background estimate.

10 Results

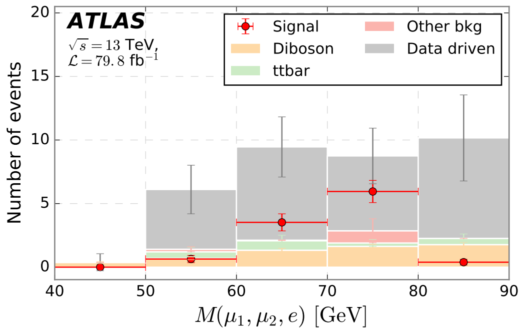
The background estimation method developed in the previous chapters is applied to the signal region and the sensitivity of the analysis is estimated. The fake background is determined using the fake factor method where a total systematic uncertainty of 23% and 22% is applied to the muon and electron fake factors respectively.

Figure 44 shows the distribution of $M(\mu_1, \mu_2, e)$ in the signal region for the five signal samples with $M_N = 5$, 10, 20, 30 and 50 GeV. Table 29 shows the corresponding significance for the five signal samples and the expected mixing angle sensitivity at 2σ based on the background-only hypothesis. The significance, see Eq.7.2, is determined using the five bins shown in Fig. 44. For the signal samples with $M_N = 5$ and 10 GeV, the estimated significance is lower than 2σ based on the background-only hypothesis. Thus for the smallest mass points, the search is sensitive to mixing angles which are larger than those used to produce the signal samples. When the mixing angle is increasing, the proper decay length for a fixed HNL mass will be smaller. For the signal samples with $M_N = 30$ and 50 GeV, the estimated significance is higher than 2σ, so the the probed mixing angles are smaller than those used to produce the samples.

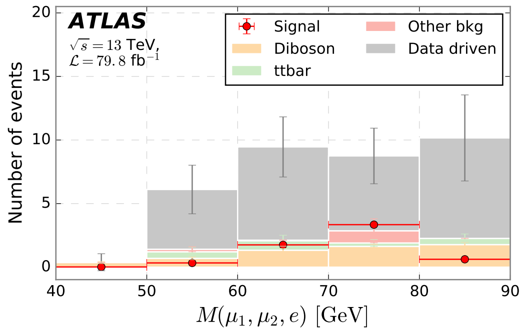
The relationship between the HNL mass, M_N, the mixing angle, $|V_{\mu N}|^2$, and the proper decay length, $c\tau$, was described in Sec. 3. It was shown that for a constant HNL mass will the proper decay length increase when then mixing angle decreases. When the mixing angle is changing should the proper decay length thus be changed accordingly. The effect on the proper decay length are thereby the signal efficiency in the analysis is not taken into account for the sensitivities listed in Tab. 29.

Figure 45 shows the best exclusion limits for the mixing angle $|V_{\mu N}|^2$ at the time of writing. In the considered region for $M_N = 5 - 50$ GeV, the best limits are currently set by the CMS [22] and DELPHI [19] experiments which lies just above $|V_{\mu N}|^2 \approx 10^{-5}$ in the considered area. Table 30 shows the sensitivity of the CMS experiment [22] for the HNL mass points applied in this analysis. The estimated sensitivity for the search in this thesis lies well below the current limits for the mass points $M_N = 20$, 30 and 50 GeV.

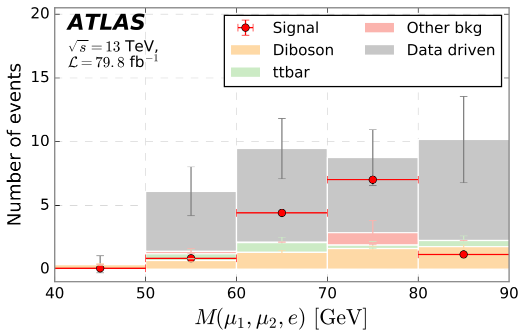
Figure 46 shows the distribution of $M(\mu_2, e)$ in the signal region for the five signal samples. A negative number of events, -0.7 ± 1.7, is estimated in the first bin of the distribution with $M(\mu_2, e) < 10$ GeV. This can simply be a statistical fluctuation but it can also be related to a bad



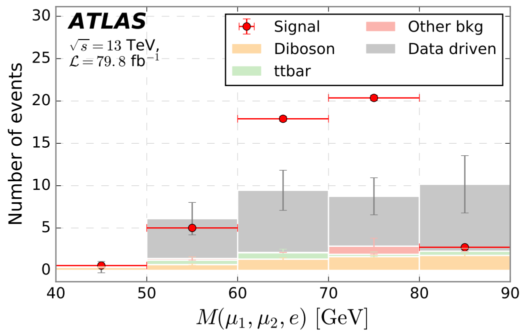
(a) Signal sample: $M_N = 5$ GeV, $c\tau = 1$ mm.



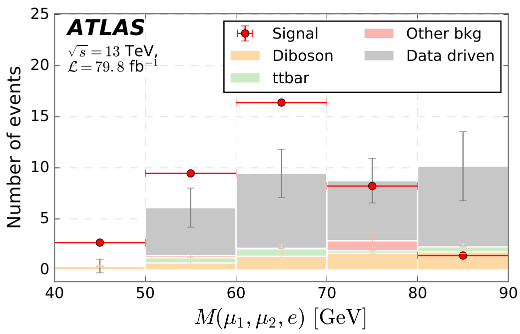
(b) Signal sample: $M_N = 10$ GeV, $c\tau = 1$ mm.



(c) Signal sample: $M_N = 20$ GeV, $c\tau = 0.1$ mm.



(d) Signal sample: $M_N = 30$ GeV, $c\tau = 10$ μm.



(e) Signal sample: $M_N = 50$ GeV, $c\tau = 1$ μm.

Figure 44: Signal region: Distribution of $M(\mu_1, \mu_2, e)$ for the five signal samples.

	Expected sensitivity of the HNL signal Based on the data-driven model				
	$M = 5$ GeV	$M = 10$ GeV	$M = 20$ GeV	$M = 30$ GeV	$M = 50$ GeV
Mixing angle	$3.9 \cdot 10^{-4}$	$1.2 \cdot 10^{-5}$	$3.8 \cdot 10^{-6}$	$4.6 \cdot 10^{-6}$	$3.4 \cdot 10^{-6}$
Significance Z	1.62	0.93	1.92	5.35	4.60
Mixing angle:	$4.9 \cdot 10^{-4}$	$2.8 \cdot 10^{-5}$	$4.0 \cdot 10^{-6}$	$1.4 \cdot 10^{-6}$	$1.3 \cdot 10^{-6}$
Significance Z	2	2	2	2	2

Table 29: The expected sensitivity of the HNL signal to the mixing angle $|V_{\mu N}|^2$. The fake background is estimated using the data-driven model developed in this analysis. All the signal region requirements listed in Tab. 11 are applied. The top rows show the significance corresponding to the the mixing angles applied in the signal samples which were determined for the specific M_N and $c\tau$ values. The bottom rows show the expected sensitivity of the search when requiring a 2σ significance for the background-only hypothesis. The significance is determined using Eq. 7.2 and binned in the $M(\mu_1, \mu_2, e)$ distribution using $N_{\text{bins}} = 5$ in the range from 40 GeV $< M(\mu_1, \mu_2, e) <$ 90 GeV.

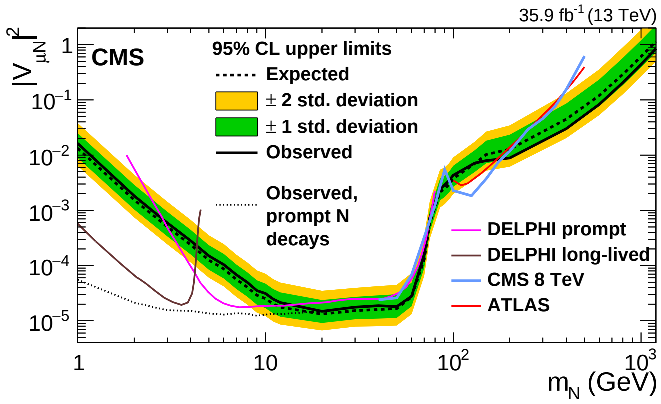


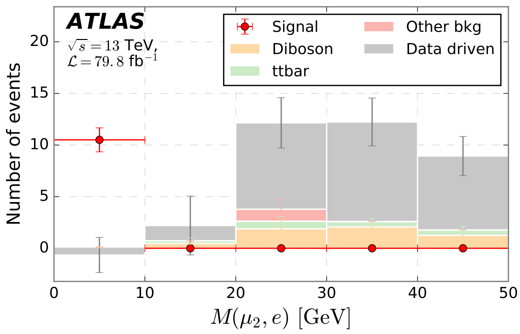
Figure 45: The exclusion region at 95% confidence level (CL) in the $|V_{\mu N}|^2$ vs M_N plane from the HNL search by the CMS collaboration [22] with 35.9 fb^{-1} at $\sqrt{s} = 13$ TeV. The best upper limits at 95% CL from the searches in DELPHI [19], ATLAS [20] at $\sqrt{s} = 8$ TeV; and CMS [21] at $\sqrt{s} = 8$ TeV. The figure is from [22].

multi-fake modeling when the spatial separation between the two leptons is small. In Sec. G, the multi-fake background estimation is studied for the cases where $\Delta R(\mu_2, e)$ is small. It is shown that the signal region requirement $\Delta R(\mu_2, e) > 0.2$ helps ensuring that the two fake leptons are independent.

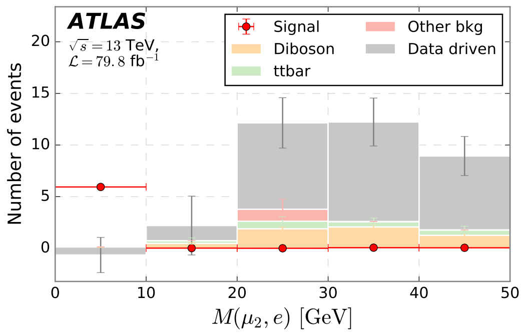
The correlation between the $M(\mu_2, e)$ distribution and the HNL mass, $M(\mu_2, e) < M_N$, can be exploited by binning the significance in the $M(\mu_2, e)$ variable. Due to the vanishing number of signal events in the first bin of the $M(\mu_2, e)$ distribution, it is not possible to get a good estimate of the sensitivity for the small signal samples using a simple counting-experiment significance, see Eq. 7.2. In order to get a proper estimate for the sensitivity of the search, it is necessary to apply an exclusion fit in the variable distribution.

	CMS limits for $\|V_{\mu N}\|^2$				
	$M = 5$ GeV	$M = 10$ GeV	$M = 20$ GeV	$M = 30$ GeV	$M = 50$ GeV
Mixing angle	$1.50 \cdot 10^{-4}$	$3.12 \cdot 10^{-5}$	$1.49 \cdot 10^{-5}$	$1.77 \cdot 10^{-5}$	$1.80 \cdot 10^{-5}$

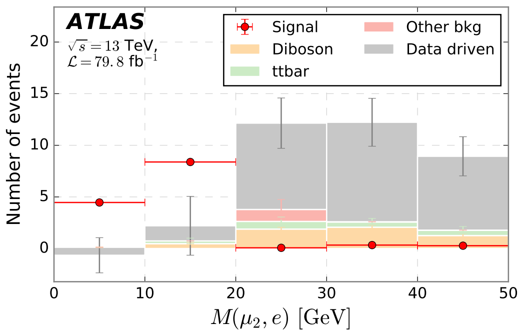
Table 30: The exclusion limits probed by the CMS experiment [22] for each of the five mass points M_N considered in this analysis.



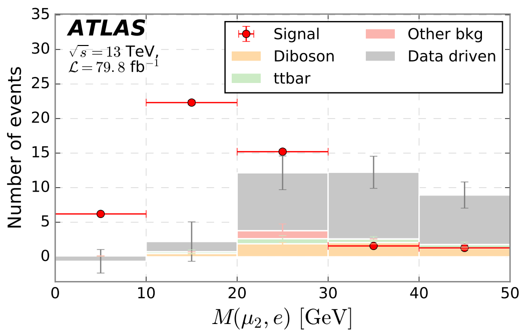
(a) Signal sample: $M_N = 5$ GeV, $c\tau = 1$ mm.



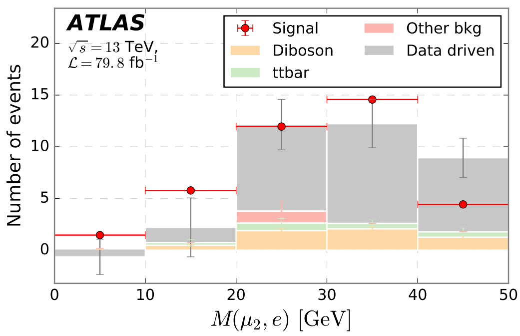
(b) Signal sample: $M_N = 10$ GeV, $c\tau = 1$ mm.



(c) Signal sample: $M_N = 20$ GeV, $c\tau = 0.1$ mm.



(d) Signal sample: $M_N = 30$ GeV, $c\tau = 10$ μm.



(e) Signal sample: $M_N = 50$ GeV, $c\tau = 1$ μm.

Figure 46: Signal region: Distribution of $M(\mu_2, e)$ for the five signal samples.

11 Conclusion

Properties of the νMSM have been examined and an analysis of HNL signals with mass-points below the W-boson mass has been performed using data recorded by the ATLAS experiment. The analysed data corresponding to 79.8 fb^{-1} is provided by the ATLAS experiment which was measured at a centre-of-mass energy of $\sqrt{s} = 13$ TeV. This is the first analysis which shows preliminary results for the search of right-handed neutrinos in the prompt trilepton channel using the full dataset from 2015-2017 recorded by the ATLAS detector. The two main parts of the analysis are: a signal selection determining the optimal requirements for the search of HNL signals in the channel $W^{\pm} \to \mu^{\pm}\mu^{\pm}e^{\mp}\nu_e$; and a background estimation based on a data-driven model using the fake factor method.

Optimizing the signal selection means maximizing the expected discovery significance for the signal where the significance is determined by a Poisson counting experiment. The trilepton final state does not lead to a sharp peak in the invariant mass spectrum due to the SM neutrino in the final state; it is therefore necessary to use different kinematic properties of the HNL events in the selection to focus the search. The discriminating variables, which are used to separate the interesting signal events from the background events, include: $M(\mu_1, \mu_2, e)$ which peaks just below the W-boson mass depending on the energy of the final state SM neutrino; $M(\mu_2, e)$ which is correlated with the mass of the HNL; E_T^{miss} which is relatively low in the signal final state compared to some background processes; and a b-jet veto. In order to maintain a high signal efficiency, it is necessary to apply loose identification and isolation selection criteria for the two leptons which decay from the HNL. This is due to the decay length of the HNL which for some of the signal samples leads to a slightly displaced vertex.

The majority of the background events in the signal region have a fake or non-prompt lepton contributing to the final state. This is due to the rare final state topology of the signal region and the relatively loose selection criteria applied for the subleading muon and electron in the signal selection. This makes it crucial to obtain a good fake background estimate for which a data-driven model is applied. The fake factor method is used to model the fake background for which fake refers to contributions from both fake leptons and real leptons arising from non-prompt decay channels. Since the leading muon in the signal finale state is almost always prompt, the fake factor method governs only the fake contribution from the subleading muon and electron. The stability of the model is tested in control regions which are similar but independent of the signal region. In these regions, the predicted background is consistent with data.

The sensitivity of the ATLAS experiment to the mixing angle $|V_{\mu N}|^2$ is estimated for signal samples with HNL masses of $M_N = 5$, 10, 20, 30 and 50 GeV. It is found that the sensitivity to the mixing angle $|V_{\mu N}|^2$ for $M_N = 20$, 30 and 50 GeV is better than the current state-of-the-art exclusion limits set by the CMS [22] and DELPHI [19] experiments. For the lowest mass point $M_N = 5$ GeV is the sensitivity lower than the current limits and for the mass point $M_N = 10$ GeV are they comparable. Based on the analysis made in this thesis, it is thus expected that the ATLAS experiment can set new improved limits to $|V_{\mu N}|^2$ which are well below the current limits in the HNL mass range from 20-50 GeV.

References

[1] M. Thomson. *Modern Particle Physics.* Cambridge University Pres, 2013.

[2] Takehiko Asaka and Mikhail Shaposhnikov. "The nuMSM, dark matter and baryon asymmetry of the universe." In: *Phys. Lett.* B620 (2005), pp. 17–26. DOI: 10.1016/j.physletb.2005.06.020. arXiv: hep-ph/0505013 [hep-ph].

[3] Mikhail Shaposhnikov and Igor Tkachev. "The nuMSM, inflation, and dark matter." In: *Phys. Lett.* B639 (2006), pp. 414–417. DOI: 10.1016/j.physletb.2006.06.063. arXiv: hep-ph/0604236 [hep-ph].

[4] Alexey Boyarsky, Oleg Ruchayskiy, and Mikhail Shaposhnikov. “The Role of sterile neutrinos in cosmology and astrophysics.” In: *Ann. Rev. Nucl. Part. Sci.* 59 (2009), pp. 191–214. DOI: 10.1146/annurev.nucl.010909.083654. arXiv: 0901.0011 [hep-ph].

[5] Oleg Ruchayskiy and Artem Ivashko. “Experimental bounds on sterile neutrino mixing angles.” In: *JHEP* 06 (2012), p. 100. DOI: 10.1007/JHEP06(2012)100. arXiv: 1112.3319 [hep-ph].

[6] Frank F. Deppisch, P. S. Bhupal Dev, and Apostolos Pilaftsis. “Neutrinos and Collider Physics.” In: *New J. Phys.* 17.7 (2015), p. 075019. DOI: 10.1088/1367-2630/17/7/075019. arXiv: 1502.06541 [hep-ph].

[7] CERN. *Home Page*. URL: https://home.cern/.

[8] ATLAS Collaboration. “Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC.” In: *Phys. Lett. B* 716 (2012), p. 1. DOI: 10.1016/j.physletb.2012.08.020. arXiv: 1207.7214 [hep-ex].

[9] CMS Collaboration. “Observation of a new boson with mass near 125 GeV in *pp* collisions at $\sqrt{s} = 7$ and 8 TeV.” In: *JHEP* 06 (2013), p. 081. DOI: 10.1007/JHEP06(2013)081. arXiv: 1303.4571 [hep-ex].

[10] A. Strumia and F. Vissani. “Neutrino masses and mixings and...” In: *ArXiv High Energy Physics* (2006). eprint: hep-ph/0606054.

[11] M. Tanabashi et al. “Review of Particle Physics.” In: *Phys. Rev. D* 98 (3 2018), p. 030001. DOI: 10.1103/PhysRevD.98.030001. URL: https://link.aps.org/doi/10.1103/PhysRevD.98.030001.

[12] B. Gripaios. *Gauge Field Theory.* Lecture notes: Gauge Field Theory at University of Cambridge Part III Natural Sciences Tripos. 2017.

[13] Eder Izaguirre and Brian Shuve. “Multilepton and Lepton Jet Probes of Sub-Weak-Scale Right-Handed Neutrinos.” In: *Phys. Rev.* D91.9 (2015), p. 093010. DOI: 10.1103/PhysRevD.91.093010. arXiv: 1504.02470 [hep-ph].

[14] Dmitry Gorbunov and Mikhail Shaposhnikov. “How to find neutral leptons of the νMSM?” In: *JHEP* 10 (2007). [Erratum: JHEP11,101(2013)], p. 015. DOI: 10.1007/JHEP11(2013)101,10.1088/1126-6708/2007/10/015. arXiv: 0705.1729 [hep-ph].

[15] ATLAS Collaboration. *ATLAS TWiki: HNL Prompt Trilepton Low Mass.* URL: https://twiki.cern.ch/twiki/bin/view/AtlasProtected/HNLPromptTriLeptonLowMass.

[16] M. C. Gonzalez-Garcia, A. Santamaria, and J. W. F. Valle. “Isosinglet Neutral Heavy Lepton Production in *Z* Decays and Neutrino Mass.” In: *Nucl. Phys.* B342 (1990), pp. 108–126. DOI: 10.1016/0550-3213(90)90573-V.

[17] Kyrylo Bondarenko et al. “Phenomenology of GeV-scale Heavy Neutral Leptons.” en. In: *arXiv:1805.08567 [hep-ex, physics:hep-ph]* (May 2018). arXiv: 1805.08567. URL: http://arxiv.org/abs/1805.08567 (visited on 07/17/2018).

[18] Biswal et al. *Search for heavy neutral leptons produced in W boson decays using a displaced-vertex signature with the ATLAS detector.* ATLAS note: EXOTICS-2017-701. 2018.

[19] P. Abreu et al. “Search for neutral heavy leptons produced in Z decays.” In: *Z. Phys.* C74 (1997). [Erratum: Z. Phys.C75,580(1997)], pp. 57–71. DOI: 10.1007/s002880050370.

[20] Georges Aad et al. “Search for heavy Majorana neutrinos with the ATLAS detector in pp collisions at $\sqrt{s} = 8$ TeV.” In: *JHEP* 07 (2015), p. 162. DOI: 10.1007/JHEP07(2015)162. arXiv: 1506.06020 [hep-ex].

[21] Vardan Khachatryan et al. “Search for heavy Majorana neutrinos in $\mu^{\pm}\mu^{\pm}+$ jets events in proton-proton collisions at $\sqrt{s} = 8$ TeV.” In: *Phys. Lett.* B748 (2015), pp. 144–166. DOI: 10.1016/j.physletb.2015.06.070. arXiv: 1501.05566 [hep-ex].

[22] Albert M Sirunyan et al. “Search for heavy neutral leptons in events with three charged leptons in proton-proton collisions at $\sqrt{s} = 13$ TeV.” In: *Phys. Rev. Lett.* 120.22 (2018), p. 221801. DOI: 10.1103/PhysRevLett.120.221801. arXiv: 1802.02965 [hep-ex].

[23] O.S. Brüning et al. (eds.) *The LHC design report 1: The LHC Main Ring.* CERN-2004-003-V-1. cdsweb.cern.ch/record/782076. 2004.

[24] O.S. Brüning et al. (eds.) *The LHC design report 2: The LHC infrastructure and general services.* CERN-2004-003-V-2. cdsweb.cern.ch/record/815187. 2004.

[25] M. Benedikt et al. (eds.) *The LHC design report 3: The LHC injector chain.* CERN-2004-003-V-3. cdsweb.cern.ch/record/823808. 2004.

[26] ATLAS Collaboration. "The ATLAS Experiment at the CERN Large Hadron Collider." In: *JINST* 3 (2008), S08003. DOI: 10.1088/1748-0221/3/08/S08003.

[27] ATLAS Collaboration. *ATLAS TWiki: Luminosity Public Results Run 2.* URL: https://twiki.cern.ch/twiki/bin/view/AtlasPublic/LuminosityPublicResultsRun2.

[28] ATLAS Collaboration. *Performance of the ATLAS Inner Detector Track and Vertex Reconstruction in High Pile-Up LHC Environment.* ATLAS-CONF-2012-042. 2012. URL: https://cds.cern.ch/record/1435196.

[29] ATLAS Collaboration. *Photos: Full detector.* URL: http://atlasexperiment.org/photos/full-detector-cgi.html.

[30] ATLAS Collaboration. *ATLAS Insertable B-Layer Technical Desgin Report.* ATLAS-TDR-19. URL: https://cds.cern.ch/record/1291633.

[31] ATLAS Collaboration. *ATLAS Insertable B-Layer Technical Design Report Addendum.* ATLAS-TDR-19-ADD-1. URL: https://cds.cern.ch/record/1451888.

[32] ATLAS Collaboration. "Performance of the ATLAS Trigger System in 2015." In: (2016). arXiv: 1611.09661 [hep-ex].

[33] ATLAS Collaboration. *Photos: How ATLAS Works.* URL: http://atlasexperiment.org/photos/how-atlas-works.html.

[34] ATLAS Collaboration. "Topological cell clustering in the ATLAS calorimeters and its performance in LHC Run 1." In: (2016). arXiv: 1603.02934 [hep-ex].

[35] ATLAS Collaboration. *ATLAS TWiki: Luminosity For Physics.* URL: https://twiki.cern.ch/twiki/bin/viewauth/Atlas/LuminosityForPhysics.

[36] ATLAS Collaboration. "Luminosity determination in pp collisions at $\sqrt{s} = 8$ TeV using the ATLAS detector at the LHC." In: *Eur. Phys. J. C* 76 (2016), p. 653. DOI: 10.1140/epjc/s10052-016-4466-1. arXiv: 1608.03953 [hep-ex].

[37] ATLAS Collaboration. *ATLAS TWiki: Good Run Lists for Analysis Run 2.* URL: https://twiki.cern.ch/twiki/bin/view/AtlasProtected/GoodRunListsForAnalysisRun2.

[38] ATLAS Collaboration. "The ATLAS Simulation Infrastructure." In: *Eur. Phys. J. C* 70 (2010), p. 823. DOI: 10.1140/epjc/s10052-010-1429-9. arXiv: 1005.4568 [hep-ex].

[39] Torbjörn Sjöstrand et al. "An Introduction to PYTHIA 8.2." In: *Comput. Phys. Commun.* 191 (2015), pp. 159–177. DOI: 10.1016/j.cpc.2015.01.024. arXiv: 1410.3012 [hep-ph].

[40] Carlo Oleari. "The POWHEG-BOX." In: *Nucl. Phys. Proc. Suppl.* 205-206 (2010), pp. 36–41. DOI: 10.1016/j.nuclphysbps.2010.08.016. arXiv: 1007.3893 [hep-ph].

[41] Johan Alwall et al. "MadGraph 5 : Going Beyond." In: *JHEP* 06 (2011), p. 128. DOI: 10.1007/JHEP06(2011)128. arXiv: 1106.0522 [hep-ph].

[42] T. Gleisberg et al. "Event generation with SHERPA 1.1." In: *JHEP* 02 (2009), p. 007. DOI: 10.1088/1126-6708/2009/02/007. arXiv: 0811.4622 [hep-ph].

[43] Sherpa Team. *Sherpa Home Page.* URL: https://sherpa.hepforge.org/trac/wiki.

[44] NNPDF Collaboration. *Home Page.* URL: http://nnpdf.mi.infn.it/.

[45] ATLAS Collaboration. *Electron efficiency measurements with the ATLAS detector using the 2012 LHC proton–proton collision data.* ATLAS-CONF-2014-032. 2014. URL: https://cds.cern.ch/record/1706245.

[46] ATLAS Collaboration. *ATLAS TWiki: Lowest Uncprescaled Triggers.* URL: https://twiki.cern.ch/twiki/bin/viewauth/Atlas/LowestUnprescaled.

[47] ATLAS Collaboration. *Electron efficiency measurements with the ATLAS detector using the 2015 LHC proton–proton collision data.* ATLAS-CONF-2016-024. 2016. URL: https://cds.cern.ch/record/2157687.

[48] ATLAS Collaboration. *Muon reconstruction performance in early $\sqrt{s} = 13$ TeV data.* ATL-PHYS-PUB-2015-037. 2015. URL: https://cds.cern.ch/record/2047831.

[49] ATLAS Collaboration. *ATLAS TWiki: Muon Selection Tool: Quality definition.* URL: https://twiki.cern.ch/twiki/bin/view/Atlas/MuonSelectionTool#Quality_definition.

[50] ATLAS Collaboration. *Jet Calibration and Systematic Uncertainties for Jets Reconstructed in the ATLAS Detector at $\sqrt{s} = 13$ TeV.* ATL-PHYS-PUB-2015-015. 2015. URL: https://cds.cern.ch/record/2037613.

[51] Matteo Cacciari, Gavin P. Salam, and Gregory Soyez. "The anti-k_t jet clustering algorithm." In: *JHEP* 04 (2008), p. 063. DOI: 10.1088/1126-6708/2008/04/063. arXiv: 0802.1189 [hep-ph].

[52] ATLAS Collaboration. *Tagging and suppression of pileup jets.* ATL-PHYS-PUB-2014-001. 2014. URL: https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/PUBNOTES/ATL-PHYS-PUB-2014-001/.

[53] ATLAS Collaboration. "Measurements of b-jet tagging efficiency with the ATLAS detector using $t\bar{t}$ events at $\sqrt{s} = 13$ TeV." In: *JHEP* 08 (2018), p. 089. DOI: 10.1007/JHEP08(2018)089. arXiv: 1805.01845 [hep-ex].

[54] ATLAS Collaboration. "Performance of missing transverse momentum reconstruction with the ATLAS detector using proton-proton collisions at $\sqrt{s} = 13$ TeV." In: (2018). arXiv: 1802.08168 [hep-ex].

[55] ATLAS Collaboration. *ATLAS TWiki: Isolation Selection Tool.* URL: https://twiki.cern.ch/twiki/bin/view/AtlasProtected/IsolationSelectionTool.

[56] ATLAS Collaboration. *ATLAS TWiki: Inner Detector Tracking Analysis Guidelines: Impact parameter definitions.* URL: https://twiki.cern.ch/twiki/bin/view/AtlasProtected/InDetTrackingDC14#Impact_parameters_z0_d0_definiti.

[57] ATLAS Collaboration. *ATLAS TWiki: Latest Recommendations for Electron ID Run 2.* URL: https://twiki.cern.ch/twiki/bin/view/AtlasProtected/LatestRecommendationsElectronIDRun2.

[58] ATLAS Collaboration. *ATLAS TWiki: Muon Combined Performance Group - Analysis Guidelines MC16.* URL: https://twiki.cern.ch/twiki/bin/view/AtlasProtected/MCPAnalysisGuidelinesMC16.

[59] Geant4. *Geant 4 - A simulation toolkit.* URL: https://geant4.web.cern.ch/.

[60] ATLAS Collaboration. *ATLAS TWiki: MCTruthClassificer Tool.* URL: https://twiki.cern.ch/twiki/bin/view/AtlasProtected/MCTruthClassifier.

[61] G. Cowan. *Discovery sensitivity for a counting experiment with background uncertainty.* 2012. URL: {https://www.pp.rhul.ac.uk/\~cowan/stat/medsig/medsigNote.pdf}.

[62] Brendlinger et al. *Overview of the methods used to estimate the fake lepton background in the SUSY group.* https://cds.cern.ch/record/2261709/files/ATL-COM-PHYS-2017-469.pdf?version=2. 2017.

[63] Abraham et al. *Search for supersymmetry with two and three leptons and missing transverse momentum in the final state at $\sqrt{s} = 13$ TeV with the ATLAS detector.* ATL-COM-PHYS-2016-1673. 2017. URL: https://cds.cern.ch/record/2233683.

[64] ATLAS Collaboration. *ATLAS TWiki: MC16 Campaign Overview.* URL: https://twiki.cern.ch/twiki/bin/view/AtlasProtected/MC16CampaignOverview.

[65] ATLAS Collaboration. *ATLAS TWiki: Derivation Production Team.* URL: https://twiki.cern.ch/twiki/bin/view/AtlasProtected/DerivationProductionTeam.

A Detailed MC sample information

Details about the MC samples applied in the analysis are provided in this section.

Within the ATLAS collaboration, the MC production is divided into so-called campaigns [64]. In this analysis, the MC16a sub-campaign is used to model the MC samples corresponding to the data period from 2015-2016 and the MC16d sub-campaign for the 2017 data. The MC16a and MC16d samples are designed to match the data conditions, trigger menu and mu-profile for the two data taking periods. The mu-profile describes the pile-up conditions in the measured data based on the number of interactions per bunch crossing. Both campaigns were launched after the data taking was complete for the corresponding period so the actual measured mu-profile is applied to get an accurate description of data. Below, the tables describing the MC sample are provided for the MC16a and MC16d sub-campaigns separately.

In Tab. 31, an overview of the applied MC samples is shown together with the MC generator used to produce the samples, the set of Parton Distribution Functions (PDF) used, and a reference to the table showing the details of the MC samples.

In addition to the name and DSID of each MC sample, the tables include the following information:

$\sigma \cdot \epsilon$	:	The cross-section provided by the MC generator multiplied with a possible truth-level filter efficiency.
k-factor	:	The k-factor is used to normalize the generator cross-section to the best known for a process.
$\sum w$	:	Sum of the generator-level weights and the pile-up reweighting effect.
$\mathcal{L}_{int}$	:	The generated equivalent integrated luminosity for the sample.

Process	**MC generator**	**PDF**	**Appendix**
Signal	Pythia v8.2	NNPDF2.3LO	Tab. 32-33
$t\bar{t}$	Powheg + Pythia v8.2	NNPDF2.3LO	Tab. 34-35
$t\bar{t}W$	MadGraph5 v2.3 + Pythia v8.2	NNPDF2.3LO	Tab. 36-37
$t\bar{t}Z$	MadGraph5 v2.2 + Pythia v8.1	NNPDF2.3LO	Tab. 36-37
t and Wt	Powheg + Pythia v6.4	Perugia2012	Tab. 38-41
W+jets	Sherpa v2.2	NNPDF3.0NLO	Tab. 42-47
Z+jets	Sherpa v2.2	NNPDF3.0NLO	Tab. 48-53
$V\gamma^*$	Sherpa v2.2	NNPDF3.0NLO	Tab. 54-55
VV: fully leptonic	Powheg + Pythia v8.2	AZNLO CTEQ6L1	Tab. 56-61
VV: partial hadronic	Sherpa v2.2	NNPDF3.0NLO	Tab. 56-61
VVV: fully leptonic	Sherpa v2.2	NNPDF3.0NLO	Tab. 62-63
VVV: partial hadronic	Powheg + Pythia v8.1	AZNLO CTEG6L1	Tab. 62-63

Table 31: List of all the MC processes applied in the analysis and the Monte Carlo (MC) generators and parton distribution functions (PDF) applied for the processes.

DSID	**Name**	$\sigma \cdot \epsilon$ [pb]	k-factor	N_{gen}	N_{all}	$\Sigma\, w$	$\mathcal{L}$ [fb^{-1}]
308986	Pythia8EvtGen_A14NNPDF23LO_WmuHNL50_5G_lt1dd_lepfilt_ch1	7.63×10^{-1}	0.24	34 596	50 000	5.00×10^{4}	2.76×10^{2}
308990	Pythia8EvtGen_A14NNPDF23LO_WmuHNL50_10G_lt1dd_lepfilt_ch1	2.39×10^{-2}	0.24	35 515	50 000	5.00×10^{4}	8.81×10^{3}
308974	Pythia8EvtGen_A14NNPDF23LO_WmuHNL50_20G_lt01dd_lepfilt_ch1	7.23×10^{-3}	0.25	35 764	50 000	5.00×10^{4}	2.71×10^{4}
999991	Pythia8EvtGen_A14NNPDF23LO_WmuHNL50_30G_lt001dd_lepfilt_ch1	9.28×10^{-3}	0.30	24 876	50 000	2.49×10^{4}	8.97×10^{3}
999994	Pythia8EvtGen_A14NNPDF23LO_WmuHNL50_50G_lt0001dd_lepfilt_ch1	6.37×10^{-3}	0.37	25 006	50 000	2.50×10^{4}	1.05×10^{4}

Table 32: MC16 sub-campaign: MC16a. Summary of signal samples used in the analysis.

DSID	**Name**	$\sigma \cdot \epsilon$ [pb]	k-factor	N_{gen}	N_{all}	$\Sigma\, w$	$\mathcal{L}$ [fb^{-1}]
308986	Pythia8EvtGen_A14NNPDF23LO_WmuHNL50_5G_lt1dd_lepfilt_ch1	7.63×10^{-1}	0.24	40 518	60 000	6.00×10^{4}	3.31×10^{2}
308990	Pythia8EvtGen_A14NNPDF23LO_WmuHNL50_10G_lt1dd_lepfilt_ch1	2.39×10^{-2}	0.24	41 257	60 000	6.00×10^{4}	1.06×10^{4}
308974	Pythia8EvtGen_A14NNPDF23LO_WmuHNL50_20G_lt01dd_lepfilt_ch1	7.23×10^{-3}	0.25	41 826	60 000	6.00×10^{4}	3.26×10^{4}
999991	Pythia8EvtGen_A14NNPDF23LO_WmuHNL50_30G_lt001dd_lepfilt_ch1	9.28×10^{-3}	0.30	24 876	60 000	2.49×10^{4}	8.97×10^{3}
999994	Pythia8EvtGen_A14NNPDF23LO_WmuHNL50_50G_lt0001dd_lepfilt_ch1	6.37×10^{-3}	0.37	25 006	60 000	2.50×10^{4}	1.05×10^{4}

Table 33: MC16 sub-campaign: MC16d. Summary of signal samples used in the analysis.

DSID	Name	$\sigma \cdot \epsilon$ [pb]	k-factor	N_{gen}	N_{all}	$\Sigma\, w$	$\mathcal{L}$ [fb^{-1}]
410470	PhPy8EG_A14_ttbar_hdamp258p75_nonallhad	3.97×10^{2}	1.14	12 978 373	59 998 000	4.38×10^{10}	9.69×10^{4}

Table 34: MC16 sub-campaign: MC16a. Summary of $t\bar{t}$ samples used in the analysis.

DSID	Name	$\sigma \cdot \epsilon$ [pb]	k-factor	N_{gen}	N_{all}	$\Sigma\, w$	$\mathcal{L}$ [fb^{-1}]
410470	PhPy8EG_A14_ttbar_hdamp258p75_nonallhad	3.97×10^{2}	1.14	15 823 308	74 486 000	5.44×10^{10}	1.20×10^{5}

Table 35: MC16 sub-campaign: MC16d. Summary of $t\bar{t}$ samples used in the analysis.

DSID	Name	$\sigma \cdot \epsilon$ [pb]	k-factor	N_{gen}	N_{all}	$\Sigma\, w$	$\mathcal{L}$ [fb^{-1}]
410155	aMcAtNloPythia8EvtGen_MEN30NLO_A14N23LO_ttW	5.48×10^{-1}	1.10	1 658 086	7 497 000	4.11×10^{6}	6.84×10^{3}
410111	MadGraphPythia8EvtGen_A14NNPDF23LO_ttee_Np0	8.83×10^{-3}	1.51	222 016	300 000	3.00×10^{5}	2.25×10^{4}
410112	MadGraphPythia8EvtGen_A14NNPDF23LO_ttee_Np1	1.44×10^{-2}	1.51	229 127	300 000	3.46×10^{5}	1.59×10^{4}
410113	MadGraphPythia8EvtGen_A14NNPDF23LO_ttmumu_Np0	8.82×10^{-3}	1.51	232 078	300 000	3.00×10^{5}	2.25×10^{4}
410114	MadGraphPythia8EvtGen_A14NNPDF23LO_ttmumu_Np1	1.44×10^{-2}	1.51	238 636	300 000	3.46×10^{5}	1.59×10^{4}
410115	MadGraphPythia8EvtGen_A14NNPDF23LO_tttautau_Np0	9.02×10^{-3}	1.51	77 522	299 000	2.99×10^{5}	2.20×10^{4}
410116	MadGraphPythia8EvtGen_A14NNPDF23LO_tttautau_Np1	1.46×10^{-2}	1.51	86 913	300 000	3.45×10^{5}	1.56×10^{4}

Table 36: MC16 sub-campaign: MC16a. Summary of $t\bar{t}V$ samples used in the analysis.

DSID	Name	$\sigma \cdot \epsilon$ [pb]	k-factor	N_{gen}	N_{all}	$\Sigma\, w$	$\mathcal{L}$ [fb^{-1}]
410155	aMcAtNloPythia8EvtGen_MEN30NLO_A14N23LO_ttW	5.48×10^{-1}	1.10	1 630 205	7 497 000	4.11×10^{6}	6.84×10^{3}
410111	MadGraphPythia8EvtGen_A14NNPDF23LO_ttee_Np0	8.83×10^{-3}	1.51	217 099	300 000	3.00×10^{5}	2.25×10^{4}
410112	MadGraphPythia8EvtGen_A14NNPDF23LO_ttee_Np1	1.44×10^{-2}	1.51	225 394	300 000	3.46×10^{5}	1.59×10^{4}
410113	MadGraphPythia8EvtGen_A14NNPDF23LO_ttmumu_Np0	8.82×10^{-3}	1.51	230 659	300 000	3.00×10^{5}	2.25×10^{4}
410114	MadGraphPythia8EvtGen_A14NNPDF23LO_ttmumu_Np1	1.44×10^{-2}	1.51	237 449	300 000	3.46×10^{5}	1.59×10^{4}
410115	MadGraphPythia8EvtGen_A14NNPDF23LO_tttautau_Np0	9.02×10^{-3}	1.51	75 142	300 000	3.00×10^{5}	2.20×10^{4}
410116	MadGraphPythia8EvtGen_A14NNPDF23LO_tttautau_Np1	1.46×10^{-2}	1.51	84 149	300 000	3.45×10^{5}	1.56×10^{4}

Table 37: MC16 sub-campaign: MC16d. Summary of $t\bar{t}V$ samples used in the analysis.

DSID	Name	$\sigma \cdot \epsilon$ [pb]	k-factor	N_{gen}	N_{all}	Σw	$\mathcal{L}$ [fb^{-1}]
410011	PowhegPythiaEvtGen_P2012_singletop_tchan_lept_top	4.37×10^{1}	1.01	402 146	5 000 000	0.00	4.95×10^{-6}
410012	PowhegPythiaEvtGen_P2012_singletop_tchan_lept_antitop	2.58×10^{1}	1.02	421 086	4 998 000	0.00	4.91×10^{-6}
410025	PowhegPythiaEvtGen_P2012_SingleTopSchan_noAllHad_top	2.05	1.00	94 505	1 000 000	0.00	9.96×10^{-7}
410026	PowhegPythiaEvtGen_P2012_SingleTopSchan_noAllHad_antitop	1.26	1.02	96 890	1 000 000	0.00	9.80×10^{-7}

Table 38: MC16 sub-campaign: MC16a. Summary of single top samples used in the analysis.

DSID	Name	$\sigma \cdot \epsilon$ [pb]	k-factor	N_{gen}	N_{all}	Σw	$\mathcal{L}$ [fb^{-1}]
410642	PhPy8EG_A14_tchan_lept_top	3.70×10^{1}	1.00	493 973	6 238 000	2.31×10^{8}	6.24×10^{3}
410643	PhPy8EG_A14_tchan_lept_antitop	2.22×10^{1}	1.00	515 051	6 250 000	1.39×10^{8}	6.26×10^{3}
410644	PowhegPythia8EvtGen_A14_singletop_schan_lept_top	2.03	1.00	232 737	2 500 000	5.07×10^{6}	2.50×10^{3}
410645	PowhegPythia8EvtGen_A14_singletop_schan_lept_antitop	1.27	1.00	240 723	2 500 000	3.17×10^{6}	2.50×10^{3}

Table 39: MC16 sub-campaign: MC16d. Summary of single top samples used in the analysis.

DSID	Name	$\sigma \cdot \epsilon$ [pb]	k-factor	N_{gen}	N_{all}	Σw	$\mathcal{L}$ [fb^{-1}]
410013	PowhegPythiaEvtGen_P2012_Wt_inclusive_top	3.40×10^{1}	1.05	513 675	5 000 000	5.00×10^{6}	1.39×10^{2}
410014	PowhegPythiaEvtGen_P2012_Wt_inclusive_antitop	3.40×10^{1}	1.05	509 616	4 968 000	4.97×10^{6}	1.39×10^{2}

Table 40: MC16 sub-campaign: MC16a. Summary of Wt samples used in the analysis.

DSID	Name	$\sigma \cdot \epsilon$ [pb]	k-factor	N_{gen}	N_{all}	Σw	$\mathcal{L}$ [fb^{-1}]
410013	PowhegPythiaEvtGen_P2012_Wt_inclusive_top	3.40×10^{1}	1.05	508 132	5 000 000	5.00×10^{6}	1.39×10^{2}
410014	PowhegPythiaEvtGen_P2012_Wt_inclusive_antitop	3.40×10^{1}	1.05	508 139	4 998 000	5.00×10^{6}	1.40×10^{2}

Table 41: MC16 sub-campaign: MC16d. Summary of Wt samples used in the analysis.

DSID	Name	$\sigma \cdot \epsilon$ [pb]	k-factor	N_{gen}	N_{all}	$\Sigma\, w$	$\mathcal{L}$ [fb^{-1}]
364170	Sherpa_221_NNPDF30NNLO_Wenu_MAXHTPTV0_70_CVetoBVeto	1.58×10^4	0.97	275 105	24 997 000	1.68×10^7	1.10
364171	Sherpa_221_NNPDF30NNLO_Wenu_MAXHTPTV0_70_CFilterBVeto	2.51×10^3	0.97	157 533	9 910 500	5.68×10^6	2.34
364172	Sherpa_221_NNPDF30NNLO_Wenu_MAXHTPTV0_70_BFilter	8.58×10^2	0.97	265 881	17 482 400	1.06×10^7	1.27×10^1
364173	Sherpa_221_NNPDF30NNLO_Wenu_MAXHTPTV70_140_CVetoBVeto	6.38×10^2	0.97	410 072	14 815 500	5.42×10^6	8.75
364174	Sherpa_221_NNPDF30NNLO_Wenu_MAXHTPTV70_140_CFilterBVeto	2.31×10^2	0.97	384 715	9 940 400	3.76×10^6	1.68×10^1
364175	Sherpa_221_NNPDF30NNLO_Wenu_MAXHTPTV70_140_BFilter	9.78×10^1	0.97	408 734	9 895 900	4.02×10^6	4.23×10^1
364176	Sherpa_221_NNPDF30NNLO_Wenu_MAXHTPTV140_280_CVetoBVeto	2.03×10^2	0.97	395 456	10 000 000	6.23×10^6	3.16×10^1
364177	Sherpa_221_NNPDF30NNLO_Wenu_MAXHTPTV140_280_CFilterBVeto	9.92×10^1	0.97	428 313	7 498 000	5.33×10^6	5.53×10^1
364178	Sherpa_221_NNPDF30NNLO_Wenu_MAXHTPTV140_280_BFilter	3.70×10^1	0.97	1 700 697	24 860 800	1.84×10^7	5.13×10^2
364179	Sherpa_221_NNPDF30NNLO_Wenu_MAXHTPTV280_500_CVetoBVeto	3.95×10^1	0.97	285 288	4 998 800	4.38×10^6	1.14×10^2
364180	Sherpa_221_NNPDF30NNLO_Wenu_MAXHTPTV280_500_CFilterBVeto	2.31×10^1	0.97	241 211	2 999 400	2.81×10^6	1.26×10^2
364181	Sherpa_221_NNPDF30NNLO_Wenu_MAXHTPTV280_500_BFilter	9.88	0.97	318 731	2 998 000	2.87×10^6	3.00×10^2
364182	Sherpa_221_NNPDF30NNLO_Wenu_MAXHTPTV500_1000	1.50×10^1	0.97	606 612	5 998 600	6.09×10^6	4.17×10^2
364183	Sherpa_221_NNPDF30NNLO_Wenu_MAXHTPTV1000_E_CMS	1.23	0.97	674 384	3 999 000	4.13×10^6	3.45×10^3

Table 42: MC16 sub-campaign: MC16a. Summary of $W \to e\nu_e$ samples used in the analysis.

DSID	Name	$\sigma \cdot \epsilon$ [pb]	k-factor	N_{gen}	N_{all}	$\Sigma\, w$	$\mathcal{L}$ [fb^{-1}]
364170	Sherpa_221_NNPDF30NNLO_Wenu_MAXHTPTV0_70_CVetoBVeto	1.58×10^4	0.97	107 793	7 619 550	5.13×10^6	3.35×10^{-1}
364171	Sherpa_221_NNPDF30NNLO_Wenu_MAXHTPTV0_70_CFilterBVeto	2.51×10^3	0.97	193 175	9 988 700	5.72×10^6	2.35
364172	Sherpa_221_NNPDF30NNLO_Wenu_MAXHTPTV0_70_BFilter	8.58×10^2	0.97	79 431	4 371 400	2.61×10^6	3.14
364173	Sherpa_221_NNPDF30NNLO_Wenu_MAXHTPTV70_140_CVetoBVeto	6.38×10^2	0.97	115 812	3 650 000	1.34×10^6	2.16
364174	Sherpa_221_NNPDF30NNLO_Wenu_MAXHTPTV70_140_CFilterBVeto	2.31×10^2	0.97	107 066	2 471 800	9.36×10^5	4.18
364175	Sherpa_221_NNPDF30NNLO_Wenu_MAXHTPTV70_140_BFilter	9.78×10^1	0.97	109 403	2 451 900	9.97×10^5	1.05×10^1
364176	Sherpa_221_NNPDF30NNLO_Wenu_MAXHTPTV140_280_CVetoBVeto	2.03×10^2	0.97	443 459	9 979 750	6.18×10^6	3.13×10^1
364177	Sherpa_221_NNPDF30NNLO_Wenu_MAXHTPTV140_280_CFilterBVeto	9.92×10^1	0.97	116 399	1 874 800	1.33×10^6	1.38×10^1
364178	Sherpa_221_NNPDF30NNLO_Wenu_MAXHTPTV140_280_BFilter	3.70×10^1	0.97	1 562 918	21 905 000	1.62×10^7	4.52×10^2
364179	Sherpa_221_NNPDF30NNLO_Wenu_MAXHTPTV280_500_CVetoBVeto	3.95×10^1	0.97	75 930	1 249 600	1.09×10^6	2.85×10^1
364180	Sherpa_221_NNPDF30NNLO_Wenu_MAXHTPTV280_500_CFilterBVeto	2.31×10^1	0.97	250 582	2 993 800	2.80×10^6	1.25×10^2
364181	Sherpa_221_NNPDF30NNLO_Wenu_MAXHTPTV280_500_BFilter	9.88	0.97	81 198	749 850	7.18×10^5	7.49×10^1
364182	Sherpa_221_NNPDF30NNLO_Wenu_MAXHTPTV500_1000	1.50×10^1	0.97	619 737	5 994 700	6.02×10^6	4.12×10^2
364183	Sherpa_221_NNPDF30NNLO_Wenu_MAXHTPTV1000_E_CMS	1.23	0.97	1 526 520	8 893 000	9.18×10^6	7.67×10^3

Table 43: MC16 sub-campaign: MC16d. Summary of $W \to e\nu_e$ samples used in the analysis.

DSID	**Name**	$\sigma \cdot \epsilon$ [pb]	k-factor	N_{gen}	N_{all}	$\Sigma\, w$	$\mathcal{L}$ [fb^{-1}]
364156	Sherpa_221_NNPDF30NNLO_Wmunu_MAXHTPTV0_70_CVetoBVeto	1.58×10^{4}	0.97	229 032	24 966 000	1.68×10^{7}	1.10
364157	Sherpa_221_NNPDF30NNLO_Wmunu_MAXHTPTV0_70_CFilterBVeto	2.51×10^{3}	0.97	137 866	9 993 000	5.73×10^{6}	2.36
364158	Sherpa_221_NNPDF30NNLO_Wmunu_MAXHTPTV0_70_BFilter	8.54×10^{2}	0.97	97 536	6 925 200	4.18×10^{6}	5.04
364159	Sherpa_221_NNPDF30NNLO_Wmunu_MAXHTPTV70_140_CVetoBVeto	6.36×10^{2}	0.97	385 895	14 964 000	5.48×10^{6}	8.88
364160	Sherpa_221_NNPDF30NNLO_Wmunu_MAXHTPTV70_140_CFilterBVeto	2.32×10^{2}	0.97	359 092	9 975 200	3.74×10^{6}	1.66×10^{1}
364161	Sherpa_221_NNPDF30NNLO_Wmunu_MAXHTPTV70_140_BFilter	7.86×10^{1}	0.97	817 038	19 915 000	8.10×10^{6}	1.06×10^{2}
364162	Sherpa_221_NNPDF30NNLO_Wmunu_MAXHTPTV140_280_CVetoBVeto	2.05×10^{2}	0.97	372 046	10 000 000	6.23×10^{6}	3.14×10^{1}
364163	Sherpa_221_NNPDF30NNLO_Wmunu_MAXHTPTV140_280_CFilterBVeto	9.92×10^{1}	0.97	404 591	7 500 000	5.33×10^{6}	5.53×10^{1}
364164	Sherpa_221_NNPDF30NNLO_Wmunu_MAXHTPTV140_280_BFilter	3.75×10^{1}	0.97	1 727 812	25 000 000	1.85×10^{7}	5.10×10^{2}
364165	Sherpa_221_NNPDF30NNLO_Wmunu_MAXHTPTV280_500_CVetoBVeto	3.95×10^{1}	0.97	274 052	4 999 000	4.38×10^{6}	1.14×10^{2}
364166	Sherpa_221_NNPDF30NNLO_Wmunu_MAXHTPTV280_500_CFilterBVeto	2.31×10^{1}	0.97	234 460	2 989 000	2.81×10^{6}	1.26×10^{2}
364167	Sherpa_221_NNPDF30NNLO_Wmunu_MAXHTPTV280_500_BFilter	9.04	0.97	321 239	2 939 500	2.82×10^{6}	3.21×10^{2}
364168	Sherpa_221_NNPDF30NNLO_Wmunu_MAXHTPTV500_1000	1.50×10^{1}	0.97	605 611	5 997 500	6.03×10^{6}	4.14×10^{2}
364169	Sherpa_221_NNPDF30NNLO_Wmunu_MAXHTPTV1000_E_CMS	1.23	0.97	692 690	3 998 000	4.11×10^{6}	3.43×10^{3}

Table 44: MC16 sub-campaign: MC16a. Summary of $W \to \mu\nu_\mu$ samples used in the analysis.

DSID	**Name**	$\sigma \cdot \epsilon$ [pb]	k-factor	N_{gen}	N_{all}	$\Sigma\, w$	$\mathcal{L}$ [fb^{-1}]
364156	Sherpa_221_NNPDF30NNLO_Wmunu_MAXHTPTV0_70_CVetoBVeto	1.58×10^{4}	0.97	351 974	23 205 950	1.56×10^{7}	1.02
364157	Sherpa_221_NNPDF30NNLO_Wmunu_MAXHTPTV0_70_CFilterBVeto	2.51×10^{3}	0.97	199 819	9 776 200	5.60×10^{6}	2.30
364158	Sherpa_221_NNPDF30NNLO_Wmunu_MAXHTPTV0_70_BFilter	8.54×10^{2}	0.97	349 143	17 330 200	1.05×10^{7}	1.26×10^{1}
364159	Sherpa_221_NNPDF30NNLO_Wmunu_MAXHTPTV70_140_CVetoBVeto	6.36×10^{2}	0.97	118 053	3 703 000	1.36×10^{6}	2.21
364160	Sherpa_221_NNPDF30NNLO_Wmunu_MAXHTPTV70_140_CFilterBVeto	2.32×10^{2}	0.97	105 945	2 486 000	9.44×10^{5}	4.20
364161	Sherpa_221_NNPDF30NNLO_Wmunu_MAXHTPTV70_140_BFilter	7.86×10^{1}	0.97	18 451	398 150	1.62×10^{5}	2.13
364162	Sherpa_221_NNPDF30NNLO_Wmunu_MAXHTPTV140_280_CVetoBVeto	2.05×10^{2}	0.97	425 051	9 863 250	6.14×10^{6}	3.09×10^{1}
364163	Sherpa_221_NNPDF30NNLO_Wmunu_MAXHTPTV140_280_CFilterBVeto	9.92×10^{1}	0.97	449 833	7 484 250	5.32×10^{6}	5.53×10^{1}
364164	Sherpa_221_NNPDF30NNLO_Wmunu_MAXHTPTV140_280_BFilter	3.75×10^{1}	0.97	14 729	200 000	1.47×10^{5}	4.04
364165	Sherpa_221_NNPDF30NNLO_Wmunu_MAXHTPTV280_500_CVetoBVeto	3.95×10^{1}	0.97	299 611	4 997 300	4.38×10^{6}	1.14×10^{2}
364166	Sherpa_221_NNPDF30NNLO_Wmunu_MAXHTPTV280_500_CFilterBVeto	2.31×10^{1}	0.97	250 030	2 998 450	2.81×10^{6}	1.25×10^{2}
364167	Sherpa_221_NNPDF30NNLO_Wmunu_MAXHTPTV280_500_BFilter	9.04	0.97	84 601	748 850	7.22×10^{5}	8.23×10^{1}
364168	Sherpa_221_NNPDF30NNLO_Wmunu_MAXHTPTV500_1000	1.50×10^{1}	0.97	629 729	5 998 500	6.03×10^{6}	4.14×10^{2}
364169	Sherpa_221_NNPDF30NNLO_Wmunu_MAXHTPTV1000_E_CMS	1.23	0.97	712 729	4 000 000	4.11×10^{6}	3.43×10^{3}

Table 45: MC16 sub-campaign: MC16d. Summary of $W \to \mu\nu_\mu$ samples used in the analysis.

DSID	Name	$\sigma \cdot \epsilon$ [pb]	k-factor	N_{gen}	N_{all}	Σw	$\mathcal{L}$ [fb^{-1}]
364184	Sherpa_221_NNPDF30NNLO_Wtaunu_MAXHTPTV0_70_CVetoBVeto	1.58×10^4	0.97	21 356	24 921 000	1.68×10^7	1.10
364185	Sherpa_221_NNPDF30NNLO_Wtaunu_MAXHTPTV0_70_CFilterBVeto	2.52×10^3	0.97	14 881	9 994 600	5.75×10^6	2.35
364186	Sherpa_221_NNPDF30NNLO_Wtaunu_MAXHTPTV0_70_BFilter	8.63×10^2	0.97	25 301	17 483 200	1.06×10^7	1.27×10^1
364187	Sherpa_221_NNPDF30NNLO_Wtaunu_MAXHTPTV70_140_CVetoBVeto	6.39×10^2	0.97	53 774	14 998 500	5.50×10^6	8.86
364188	Sherpa_221_NNPDF30NNLO_Wtaunu_MAXHTPTV70_140_CFilterBVeto	2.29×10^2	0.97	55 921	9 999 000	3.77×10^6	1.69×10^1
364189	Sherpa_221_NNPDF30NNLO_Wtaunu_MAXHTPTV70_140_BFilter	9.83×10^1	0.97	61 827	9 999 000	4.03×10^6	4.22×10^1
364190	Sherpa_221_NNPDF30NNLO_Wtaunu_MAXHTPTV140_280_CVetoBVeto	2.03×10^2	0.97	64 455	9 970 000	6.21×10^6	3.15×10^1
364191	Sherpa_221_NNPDF30NNLO_Wtaunu_MAXHTPTV140_280_CFilterBVeto	9.67×10^1	0.97	76 076	7 500 000	5.32×10^6	5.67×10^1
364193	Sherpa_221_NNPDF30NNLO_Wtaunu_MAXHTPTV280_500_CVetoBVeto	3.95×10^1	0.97	53 767	4 999 200	4.38×10^6	1.14×10^2
364194	Sherpa_221_NNPDF30NNLO_Wtaunu_MAXHTPTV280_500_CFilterBVeto	2.30×10^1	0.97	48 168	2 998 400	2.81×10^6	1.26×10^2
364195	Sherpa_221_NNPDF30NNLO_Wtaunu_MAXHTPTV280_500_BFilter	9.78	0.97	74 465	2 998 100	2.87×10^6	3.03×10^2
364196	Sherpa_221_NNPDF30NNLO_Wtaunu_MAXHTPTV500_1000	1.51×10^1	0.97	140 973	5 960 000	6.00×10^6	4.11×10^2
364197	Sherpa_221_NNPDF30NNLO_Wtaunu_MAXHTPTV1000_E_CMS	1.23	0.97	225 668	4 000 000	4.11×10^6	3.44×10^3

Table 46: MC16 sub-campaign: MC16a. Summary of $W \to \tau\nu_\tau$ samples used in the analysis.

DSID	Name	$\sigma \cdot \epsilon$ [pb]	k-factor	N_{gen}	N_{all}	Σw	$\mathcal{L}$ [fb^{-1}]
364184	Sherpa_221_NNPDF30NNLO_Wtaunu_MAXHTPTV0_70_CVetoBVeto	1.58×10^4	0.97	6694	6 239 000	4.20×10^6	2.74×10^{-1}
364185	Sherpa_221_NNPDF30NNLO_Wtaunu_MAXHTPTV0_70_CFilterBVeto	2.52×10^3	0.97	21 367	12 443 600	7.15×10^6	2.93
364186	Sherpa_221_NNPDF30NNLO_Wtaunu_MAXHTPTV0_70_BFilter	8.63×10^2	0.97	34 802	21 824 500	1.33×10^7	1.58×10^1
364187	Sherpa_221_NNPDF30NNLO_Wtaunu_MAXHTPTV70_140_CVetoBVeto	6.39×10^2	0.97	71 482	18 721 000	6.80×10^6	1.10×10^1
364188	Sherpa_221_NNPDF30NNLO_Wtaunu_MAXHTPTV70_140_CFilterBVeto	2.29×10^2	0.97	72 820	12 389 000	4.70×10^6	2.11×10^1
364189	Sherpa_221_NNPDF30NNLO_Wtaunu_MAXHTPTV70_140_BFilter	9.83×10^1	0.97	62 109	9 977 850	4.03×10^6	4.23×10^1
364190	Sherpa_221_NNPDF30NNLO_Wtaunu_MAXHTPTV140_280_CVetoBVeto	2.03×10^2	0.97	68 337	9 979 000	6.22×10^6	3.15×10^1
364191	Sherpa_221_NNPDF30NNLO_Wtaunu_MAXHTPTV140_280_CFilterBVeto	9.67×10^1	0.97	77 895	7 498 950	5.32×10^6	5.67×10^1
364192	Sherpa_221_NNPDF30NNLO_Wtaunu_MAXHTPTV140_280_BFilter	3.57×10^1	0.97	417 714	31 843 100	2.36×10^7	6.83×10^2
364193	Sherpa_221_NNPDF30NNLO_Wtaunu_MAXHTPTV280_500_CVetoBVeto	3.95×10^1	0.97	67 672	6 238 800	5.47×10^6	1.43×10^2
364194	Sherpa_221_NNPDF30NNLO_Wtaunu_MAXHTPTV280_500_CFilterBVeto	2.30×10^1	0.97	11 780	729 800	6.84×10^5	3.07×10^1
364195	Sherpa_221_NNPDF30NNLO_Wtaunu_MAXHTPTV280_500_BFilter	9.78	0.97	18 056	749 500	7.18×10^5	7.57×10^1
364196	Sherpa_221_NNPDF30NNLO_Wtaunu_MAXHTPTV500_1000	1.51×10^1	0.97	139 548	5 995 450	6.04×10^6	4.13×10^2
364197	Sherpa_221_NNPDF30NNLO_Wtaunu_MAXHTPTV1000_E_CMS	1.23	0.97	286 798	4 989 800	5.12×10^6	4.28×10^3

Table 47: MC16 sub-campaign: MC16d. Summary of $W \to \tau\nu_\tau$ samples used in the analysis.

DSID	**Name**	$\sigma \cdot \epsilon$ [pb]	k-factor	N_{gen}	N_{all}	$\Sigma\, w$	$\mathcal{L}$ [fb^{-1}]
364114	Sherpa_221_NNPDF30NNLO_Zee_MAXHTPTV0_70_CVetoBVeto	1.63×10^{3}	0.98	3 255 935	8 000 000	5.37×10^{6}	3.39
364115	Sherpa_221_NNPDF30NNLO_Zee_MAXHTPTV0_70_CFilterBVeto	2.26×10^{2}	0.98	2 303 894	4 999 000	2.87×10^{6}	1.30×10^{1}
364116	Sherpa_221_NNPDF30NNLO_Zee_MAXHTPTV0_70_BFilter	1.30×10^{2}	0.98	4 067 669	7 995 600	4.11×10^{6}	3.23×10^{1}
364117	Sherpa_221_NNPDF30NNLO_Zee_MAXHTPTV70_140_CVetoBVeto	7.68×10^{1}	0.98	3 120 322	5 956 000	2.18×10^{6}	2.91×10^{1}
364118	Sherpa_221_NNPDF30NNLO_Zee_MAXHTPTV70_140_CFilterBVeto	2.07×10^{1}	0.98	1 116 513	1 999 200	7.25×10^{5}	3.59×10^{1}
364119	Sherpa_221_NNPDF30NNLO_Zee_MAXHTPTV70_140_BFilter	1.32×10^{1}	0.98	3 575 342	5 970 000	2.08×10^{6}	1.62×10^{2}
364120	Sherpa_221_NNPDF30NNLO_Zee_MAXHTPTV140_280_CVetoBVeto	2.51×10^{1}	0.98	2 896 556	5 000 000	3.00×10^{6}	1.23×10^{2}
364121	Sherpa_221_NNPDF30NNLO_Zee_MAXHTPTV140_280_CFilterBVeto	9.47	0.98	1 794 036	3 000 000	2.00×10^{6}	2.17×10^{2}
364122	Sherpa_221_NNPDF30NNLO_Zee_MAXHTPTV140_280_BFilter	6.24	0.98	7 881 564	12 439 600	8.63×10^{6}	1.42×10^{3}
364123	Sherpa_221_NNPDF30NNLO_Zee_MAXHTPTV280_500_CVetoBVeto	4.91	0.98	1 225 462	1 999 800	1.72×10^{6}	3.60×10^{2}
364124	Sherpa_221_NNPDF30NNLO_Zee_MAXHTPTV280_500_CFilterBVeto	2.31	0.98	621 495	999 900	9.18×10^{5}	4.08×10^{2}
364125	Sherpa_221_NNPDF30NNLO_Zee_MAXHTPTV280_500_BFilter	1.53	0.98	1 313 741	1 999 850	1.88×10^{6}	1.26×10^{3}
364126	Sherpa_221_NNPDF30NNLO_Zee_MAXHTPTV500_1000	1.81	0.98	1 960 701	2 999 000	2.97×10^{6}	1.68×10^{3}
364127	Sherpa_221_NNPDF30NNLO_Zee_MAXHTPTV1000_E_CMS	1.49×10^{-1}	0.98	715 276	1 000 000	1.02×10^{6}	7.01×10^{3}

Table 48: MC16 sub-campaign: MC16a. Summary of $Z \to e^{\pm}e^{\mp}$ samples used in the analysis.

DSID	**Name**	$\sigma \cdot \epsilon$ [pb]	k-factor	N_{gen}	N_{all}	$\Sigma\, w$	$\mathcal{L}$ [fb^{-1}]
364115	Sherpa_221_NNPDF30NNLO_Zee_MAXHTPTV0_70_CFilterBVeto	2.26×10^{2}	0.98	1 939 031	4 418 500	2.54×10^{6}	1.15×10^{1}
364116	Sherpa_221_NNPDF30NNLO_Zee_MAXHTPTV0_70_BFilter	1.30×10^{2}	0.98	4 862 414	9 992 800	5.14×10^{6}	4.04×10^{1}
364117	Sherpa_221_NNPDF30NNLO_Zee_MAXHTPTV70_140_CVetoBVeto	7.68×10^{1}	0.98	3 801 284	7 384 500	2.70×10^{6}	3.61×10^{1}
364118	Sherpa_221_NNPDF30NNLO_Zee_MAXHTPTV70_140_CFilterBVeto	2.07×10^{1}	0.98	1 364 766	2 495 000	9.07×10^{5}	4.50×10^{1}
364119	Sherpa_221_NNPDF30NNLO_Zee_MAXHTPTV70_140_BFilter	1.32×10^{1}	0.98	1 206 231	2 052 600	7.20×10^{5}	5.59×10^{1}
364120	Sherpa_221_NNPDF30NNLO_Zee_MAXHTPTV140_280_CVetoBVeto	2.51×10^{1}	0.98	3 583 819	6 250 000	3.74×10^{6}	1.53×10^{2}
364121	Sherpa_221_NNPDF30NNLO_Zee_MAXHTPTV140_280_CFilterBVeto	9.47	0.98	2 214 067	3 748 000	2.50×10^{6}	2.71×10^{2}
364122	Sherpa_221_NNPDF30NNLO_Zee_MAXHTPTV140_280_BFilter	6.24	0.98	9 800 290	15 662 200	1.09×10^{7}	1.79×10^{3}
364123	Sherpa_221_NNPDF30NNLO_Zee_MAXHTPTV280_500_CVetoBVeto	4.91	0.98	1 520 041	2 499 800	2.15×10^{6}	4.49×10^{2}
364124	Sherpa_221_NNPDF30NNLO_Zee_MAXHTPTV280_500_CFilterBVeto	2.31	0.98	768 748	1 250 000	1.15×10^{6}	5.10×10^{2}
364125	Sherpa_221_NNPDF30NNLO_Zee_MAXHTPTV280_500_BFilter	1.53	0.98	1 623 987	2 499 500	2.34×10^{6}	1.57×10^{3}
364126	Sherpa_221_NNPDF30NNLO_Zee_MAXHTPTV500_1000	1.81	0.98	2 405 265	3 712 000	3.68×10^{6}	2.08×10^{3}
364127	Sherpa_221_NNPDF30NNLO_Zee_MAXHTPTV1000_E_CMS	1.49×10^{-1}	0.98	889 909	1 249 800	1.27×10^{6}	8.76×10^{3}

Table 49: MC16 sub-campaign: MC16d. Summary of $Z \to e^{\pm}e^{\mp}$ samples used in the analysis.

DSID	**Name**	$\sigma \cdot \epsilon$ [pb]	k-factor	N_{gen}	N_{all}	$\Sigma\, w$	$\mathcal{L}$ [fb^{-1}]
364100	Sherpa_221_NNPDF30NNLO_Zmumu_MAXHTPTV0_70_CVetoBVeto	1.63×10^{3}	0.98	3 711 187	7 970 000	5.37×10^{6}	3.38
364101	Sherpa_221_NNPDF30NNLO_Zmumu_MAXHTPTV0_70_CFilterBVeto	2.25×10^{2}	0.98	2 605 577	4 982 000	2.87×10^{6}	1.31×10^{1}
364102	Sherpa_221_NNPDF30NNLO_Zmumu_MAXHTPTV0_70_BFilter	1.30×10^{2}	0.98	4 569 941	7 983 000	4.12×10^{6}	3.24×10^{1}
364103	Sherpa_221_NNPDF30NNLO_Zmumu_MAXHTPTV70_140_CVetoBVeto	7.52×10^{1}	0.98	3 449 307	5 983 000	2.17×10^{6}	2.95×10^{1}
364104	Sherpa_221_NNPDF30NNLO_Zmumu_MAXHTPTV70_140_CFilterBVeto	2.13×10^{1}	0.98	1 219 839	1 996 800	7.32×10^{5}	3.52×10^{1}
364105	Sherpa_221_NNPDF30NNLO_Zmumu_MAXHTPTV70_140_BFilter	1.28×10^{1}	0.98	3 887 959	5 981 600	2.08×10^{6}	1.67×10^{2}
364106	Sherpa_221_NNPDF30NNLO_Zmumu_MAXHTPTV140_280_CVetoBVeto	2.40×10^{1}	0.98	3 124 953	5 000 000	2.97×10^{6}	1.27×10^{2}
364107	Sherpa_221_NNPDF30NNLO_Zmumu_MAXHTPTV140_280_CFilterBVeto	9.38	0.98	1 932 036	3 000 000	1.99×10^{6}	2.18×10^{2}
364108	Sherpa_221_NNPDF30NNLO_Zmumu_MAXHTPTV140_280_BFilter	6.23	0.98	8 433 235	12 458 900	8.65×10^{6}	1.42×10^{3}
364109	Sherpa_221_NNPDF30NNLO_Zmumu_MAXHTPTV280_500_CVetoBVeto	4.78	0.98	1 314 945	2 000 000	1.73×10^{6}	3.71×10^{2}
364110	Sherpa_221_NNPDF30NNLO_Zmumu_MAXHTPTV280_500_CFilterBVeto	2.27	0.98	667 636	999 600	9.19×10^{5}	4.15×10^{2}
364111	Sherpa_221_NNPDF30NNLO_Zmumu_MAXHTPTV280_500_BFilter	1.51	0.98	1 399 543	1 999 400	1.88×10^{6}	1.28×10^{3}
364112	Sherpa_221_NNPDF30NNLO_Zmumu_MAXHTPTV500_1000	1.79	0.98	2 113 128	2 996 500	2.98×10^{6}	1.71×10^{3}
364113	Sherpa_221_NNPDF30NNLO_Zmumu_MAXHTPTV1000_E_CMS	1.48×10^{-1}	0.98	781 017	997 000	1.02×10^{6}	7.07×10^{3}

Table 50: MC16 sub-campaign: MC16a. Summary of $Z \to \mu^{\pm}\mu^{\mp}$ samples used in the analysis.

DSID	**Name**	$\sigma \cdot \epsilon$ [pb]	k-factor	N_{gen}	N_{all}	$\Sigma\, w$	$\mathcal{L}$ [fb^{-1}]
363123	MGPy8EG_N30NLO_Zmumu_Ht0_70_CVetoBVeto	1.43×10^{3}	1.14	911 838	1 999 000	2.05×10^{6}	1.26
363124	MGPy8EG_N30NLO_Zmumu_Ht0_70_CFilterBVeto	1.87×10^{2}	1.14	311 367	580 000	6.02×10^{5}	2.83
363125	MGPy8EG_N30NLO_Zmumu_Ht0_70_BFilter	1.02×10^{2}	1.14	291 288	500 000	5.25×10^{5}	4.52
363126	MGPy8EG_N30NLO_Zmumu_Ht70_140_CVetoBVeto	6.08×10^{1}	1.14	1 123 814	1 986 000	2.78×10^{6}	4.00×10^{1}
363127	MGPy8EG_N30NLO_Zmumu_Ht70_140_CFilterBVeto	1.47×10^{1}	1.14	690 433	1 099 000	1.58×10^{6}	9.39×10^{1}
363128	MGPy8EG_N30NLO_Zmumu_Ht70_140_BFilter	9.19	1.14	268 242	400 000	6.06×10^{5}	5.78×10^{1}
363129	MGPy8EG_N30NLO_Zmumu_Ht140_280_CVetoBVeto	2.42×10^{1}	1.14	1 193 985	2 000 000	3.09×10^{6}	1.12×10^{2}
363130	MGPy8EG_N30NLO_Zmumu_Ht140_280_CFilterBVeto	7.19	1.14	364 595	568 000	9.03×10^{5}	1.10×10^{2}
363132	MGPy8EG_N30NLO_Zmumu_Ht280_500_CVetoBVeto	5.15	1.14	504 584	790 000	1.29×10^{6}	2.20×10^{2}
363133	MGPy8EG_N30NLO_Zmumu_Ht280_500_CFilterBVeto	1.87	1.14	266 117	400 000	6.70×10^{5}	3.14×10^{2}
363134	MGPy8EG_N30NLO_Zmumu_Ht280_500_BFilter	1.17	1.14	168 695	240 000	4.17×10^{5}	3.12×10^{2}
363135	MGPy8EG_N30NLO_Zmumu_Ht500_700_CVetoBVeto	7.58×10^{-1}	1.14	138 166	200 000	3.33×10^{5}	3.85×10^{2}
363136	MGPy8EG_N30NLO_Zmumu_Ht500_700_CFilterBVeto	3.16×10^{-1}	1.14	148 668	210 000	3.54×10^{5}	9.81×10^{2}
363137	MGPy8EG_N30NLO_Zmumu_Ht500_700_BFilter	1.95×10^{-1}	1.14	17 709	24 000	4.13×10^{4}	1.85×10^{2}
363138	MGPy8EG_N30NLO_Zmumu_Ht700_1000_CVetoBVeto	2.53×10^{-1}	1.14	3685	5000	8.39×10^{3}	2.90×10^{1}
363139	MGPy8EG_N30NLO_Zmumu_Ht700_1000_CFilterBVeto	1.16×10^{-1}	1.14	3751	5000	8.37×10^{3}	6.31×10^{1}

Table 51: MC16 sub-campaign: MC16d. Summary of $Z \to \mu^{\pm}\mu^{\mp}$ samples used in the analysis.

DSID	**Name**	$\sigma \cdot \epsilon$ [pb]	k-factor	N_{gen}	N_{all}	$\Sigma\, w$	$\mathcal{L}$ [fb^{-1}]
364128	Sherpa_221_NNPDF30NNLO_Ztautau_MAXHTPTV0_70_CVetoBVeto	1.63×10^3	0.98	128 009	7 986 000	5.38×10^6	3.39
364129	Sherpa_221_NNPDF30NNLO_Ztautau_MAXHTPTV0_70_CFilterBVeto	2.17×10^2	0.98	91 439	4 979 000	2.87×10^6	1.36×10^1
364130	Sherpa_221_NNPDF30NNLO_Ztautau_MAXHTPTV0_70_BFilter	1.30×10^2	0.98	179 125	7 995 800	4.11×10^6	3.24×10^1
364131	Sherpa_221_NNPDF30NNLO_Ztautau_MAXHTPTV70_140_CVetoBVeto	7.66×10^1	0.98	168 780	5 997 500	2.19×10^6	2.93×10^1
364132	Sherpa_221_NNPDF30NNLO_Ztautau_MAXHTPTV70_140_CFilterBVeto	2.10×10^1	0.98	62 634	1 999 200	7.32×10^5	3.57×10^1
364133	Sherpa_221_NNPDF30NNLO_Ztautau_MAXHTPTV70_140_BFilter	1.23×10^1	0.98	249 002	5 969 550	2.09×10^6	1.75×10^2
364134	Sherpa_221_NNPDF30NNLO_Ztautau_MAXHTPTV140_280_CVetoBVeto	2.52×10^1	0.98	196 856	4 650 000	2.79×10^6	1.13×10^2
364135	Sherpa_221_NNPDF30NNLO_Ztautau_MAXHTPTV140_280_CFilterBVeto	9.54	0.98	136 872	3 000 000	2.00×10^6	2.15×10^2
364136	Sherpa_221_NNPDF30NNLO_Ztautau_MAXHTPTV140_280_BFilter	6.35	0.98	308 827	4 992 950	3.47×10^6	5.61×10^2
364137	Sherpa_221_NNPDF30NNLO_Ztautau_MAXHTPTV280_500_CVetoBVeto	4.88	0.98	115 544	2 000 000	1.68×10^6	3.53×10^2
364138	Sherpa_221_NNPDF30NNLO_Ztautau_MAXHTPTV280_500_CFilterBVeto	2.30	0.98	61 384	1 000 000	9.18×10^5	4.09×10^2
364139	Sherpa_221_NNPDF30NNLO_Ztautau_MAXHTPTV280_500_BFilter	1.53	0.98	168 590	1 983 950	1.86×10^6	1.25×10^3
364140	Sherpa_221_NNPDF30NNLO_Ztautau_MAXHTPTV500_1000	1.81	0.98	244 580	2 999 800	2.98×10^6	1.69×10^3
364141	Sherpa_221_NNPDF30NNLO_Ztautau_MAXHTPTV1000_E_CMS	1.48×10^{-1}	0.98	131 253	1 000 000	1.02×10^6	7.04×10^3

Table 52: MC16 sub-campaign: MC16a. Summary of $Z \to \tau^{\pm}\tau^{\mp}$ samples used in the analysis.

DSID	**Name**	$\sigma \cdot \epsilon$ [pb]	k-factor	N_{gen}	N_{all}	$\Sigma\, w$	$\mathcal{L}$ [fb^{-1}]
364128	Sherpa_221_NNPDF30NNLO_Ztautau_MAXHTPTV0_70_CVetoBVeto	1.63×10^3	0.98	141 194	9 998 000	6.73×10^6	4.24
364129	Sherpa_221_NNPDF30NNLO_Ztautau_MAXHTPTV0_70_CFilterBVeto	2.17×10^2	0.98	100 006	6 137 500	3.53×10^6	1.67×10^1
364130	Sherpa_221_NNPDF30NNLO_Ztautau_MAXHTPTV0_70_BFilter	1.30×10^2	0.98	197 008	9 994 000	5.14×10^6	4.04×10^1
364131	Sherpa_221_NNPDF30NNLO_Ztautau_MAXHTPTV70_140_CVetoBVeto	7.66×10^1	0.98	195 201	7 498 500	2.76×10^6	3.69×10^1
364132	Sherpa_221_NNPDF30NNLO_Ztautau_MAXHTPTV70_140_CFilterBVeto	2.10×10^1	0.98	72 484	2 499 000	9.03×10^5	4.40×10^1
364133	Sherpa_221_NNPDF30NNLO_Ztautau_MAXHTPTV70_140_BFilter	1.23×10^1	0.98	288 131	7 498 800	2.63×10^6	2.20×10^2
364134	Sherpa_221_NNPDF30NNLO_Ztautau_MAXHTPTV140_280_CVetoBVeto	2.52×10^1	0.98	249 564	6 228 000	3.73×10^6	1.52×10^2
364135	Sherpa_221_NNPDF30NNLO_Ztautau_MAXHTPTV140_280_CFilterBVeto	9.54	0.98	163 417	3 749 900	2.50×10^6	2.69×10^2
364136	Sherpa_221_NNPDF30NNLO_Ztautau_MAXHTPTV140_280_BFilter	6.35	0.98	361 761	6 213 000	4.32×10^6	6.98×10^2
364137	Sherpa_221_NNPDF30NNLO_Ztautau_MAXHTPTV280_500_CVetoBVeto	4.88	0.98	138 184	2 499 000	2.16×10^6	4.53×10^2
364138	Sherpa_221_NNPDF30NNLO_Ztautau_MAXHTPTV280_500_CFilterBVeto	2.30	0.98	72 128	1 233 900	1.13×10^6	5.05×10^2
364139	Sherpa_221_NNPDF30NNLO_Ztautau_MAXHTPTV280_500_BFilter	1.53	0.98	202 654	2 497 200	2.34×10^6	1.57×10^3
364140	Sherpa_221_NNPDF30NNLO_Ztautau_MAXHTPTV500_1000	1.81	0.98	292 691	3 749 600	3.72×10^6	2.11×10^3
364141	Sherpa_221_NNPDF30NNLO_Ztautau_MAXHTPTV1000_E_CMS	1.48×10^{-1}	0.98	162 864	1 250 000	1.27×10^6	8.80×10^3

Table 53: MC16 sub-campaign: MC16d. Summary of $Z \to \tau^{\pm}\tau^{\mp}$ samples used in the analysis.

DSID	Name	$\sigma \cdot \epsilon$ [pb]	k-factor	N_{gen}	N_{all}	$\Sigma\, w$	$\mathcal{L}$ [fb^{-1}]
364521	Sherpa_222_NNPDF30NNLO_enugamma_pty_7_15	1.99×10^{2}	1.00	260 072	12 488 000	7.67×10^{6}	3.85×10^{1}
364522	Sherpa_222_NNPDF30NNLO_enugamma_pty_15_35	1.34×10^{2}	1.00	385 197	12 997 000	8.32×10^{6}	6.19×10^{1}
364523	Sherpa_222_NNPDF30NNLO_enugamma_pty_35_70	1.91×10^{1}	1.00	71 856	1 984 000	1.15×10^{6}	6.02×10^{1}
364524	Sherpa_222_NNPDF30NNLO_enugamma_pty_70_140	1.92	1.00	23 889	250 000	1.76×10^{5}	9.19×10^{1}
364525	Sherpa_222_NNPDF30NNLO_enugamma_pty_140_E_CMS	2.97×10^{-1}	1.00	14 141	98 500	8.17×10^{4}	2.75×10^{2}
364526	Sherpa_222_NNPDF30NNLO_munugamma_pty_7_15	1.99×10^{2}	1.00	285 747	11 939 000	7.33×10^{6}	3.68×10^{1}
364528	Sherpa_222_NNPDF30NNLO_munugamma_pty_35_70	1.91×10^{1}	1.00	75 641	2 000 000	1.16×10^{6}	6.08×10^{1}
364529	Sherpa_222_NNPDF30NNLO_munugamma_pty_70_140	1.92	1.00	23 463	250 000	1.76×10^{5}	9.17×10^{1}
364530	Sherpa_222_NNPDF30NNLO_munugamma_pty_140_E_CMS	2.98×10^{-1}	1.00	14 360	99 500	8.15×10^{4}	2.74×10^{2}
364531	Sherpa_222_NNPDF30NNLO_taunugamma_pty_7_15	1.99×10^{2}	1.00	22 981	12 411 000	7.62×10^{6}	3.82×10^{1}
364532	Sherpa_222_NNPDF30NNLO_taunugamma_pty_15_35	1.34×10^{2}	1.00	43 110	12 999 000	8.29×10^{6}	6.17×10^{1}
364533	Sherpa_222_NNPDF30NNLO_taunugamma_pty_35_70	1.91×10^{1}	1.00	12 371	1 971 000	1.14×10^{6}	5.98×10^{1}
364534	Sherpa_222_NNPDF30NNLO_taunugamma_pty_70_140	1.93	1.00	5120	247 500	1.75×10^{5}	9.09×10^{1}
364535	Sherpa_222_NNPDF30NNLO_taunugamma_pty_140_E_CMS	2.97×10^{-1}	1.00	4330	99 500	8.21×10^{4}	2.76×10^{2}
364500	Sherpa_222_NNPDF30NNLO_eegamma_pty_7_15	5.76×10^{1}	1.00	1 282 109	4 000 000	2.53×10^{6}	4.39×10^{1}
364501	Sherpa_222_NNPDF30NNLO_eegamma_pty_15_35	3.46×10^{1}	1.00	1 057 935	3 000 000	1.95×10^{6}	5.63×10^{1}
364502	Sherpa_222_NNPDF30NNLO_eegamma_pty_35_70	6.29	1.00	298 651	1 000 000	6.21×10^{5}	9.89×10^{1}
364503	Sherpa_222_NNPDF30NNLO_eegamma_pty_70_140	4.92×10^{-1}	1.00	238 466	487 000	3.40×10^{5}	6.91×10^{2}
364504	Sherpa_222_NNPDF30NNLO_eegamma_pty_140_E_CMS	6.30×10^{-2}	1.00	65 484	100 000	7.32×10^{4}	1.16×10^{3}
364505	Sherpa_222_NNPDF30NNLO_mumugamma_pty_7_15	5.77×10^{1}	1.00	1 515 986	4 000 000	2.53×10^{6}	4.39×10^{1}
364506	Sherpa_222_NNPDF30NNLO_mumugamma_pty_15_35	3.46×10^{1}	1.00	1 218 751	2 956 000	1.92×10^{6}	5.54×10^{1}
364507	Sherpa_222_NNPDF30NNLO_mumugamma_pty_35_70	6.29	1.00	343 036	1 000 000	6.17×10^{5}	9.82×10^{1}
364508	Sherpa_222_NNPDF30NNLO_mumugamma_pty_70_140	4.94×10^{-1}	1.00	275 274	498 500	3.47×10^{5}	7.02×10^{2}
364509	Sherpa_222_NNPDF30NNLO_mumugamma_pty_140_E_CMS	6.31×10^{-2}	1.00	71 061	100 000	7.35×10^{4}	1.16×10^{3}
364510	Sherpa_222_NNPDF30NNLO_tautaugamma_pty_7_15	5.76×10^{1}	1.00	48 504	4 000 000	2.53×10^{6}	4.39×10^{1}
364511	Sherpa_222_NNPDF30NNLO_tautaugamma_pty_15_35	3.46×10^{1}	1.00	46 631	2 862 000	1.86×10^{6}	5.37×10^{1}
364512	Sherpa_222_NNPDF30NNLO_tautaugamma_pty_35_70	6.30	1.00	20 054	996 000	6.16×10^{5}	9.78×10^{1}
364513	Sherpa_222_NNPDF30NNLO_tautaugamma_pty_70_140	4.95×10^{-1}	1.00	24 401	496 000	3.47×10^{5}	7.01×10^{2}
364514	Sherpa_222_NNPDF30NNLO_tautaugamma_pty_140_E_CMS	6.34×10^{-2}	1.00	10 009	100 000	7.37×10^{4}	1.16×10^{3}

Table 54: MC16 sub-campaign: MC16a. Summary of $V\gamma^{*}$ samples used in the analysis.

DSID	Name	$\sigma \cdot \epsilon$ [pb]	k-factor	N_{gen}	N_{all}	$\Sigma\, w$	$\mathcal{L}$ [fb^{-1}]
364521	Sherpa_222_NNPDF30NNLO_enugamma_pty_7_15	1.99×10^{2}	1.00	299 771	15 624 000	9.58×10^{6}	4.81×10^{1}
364522	Sherpa_222_NNPDF30NNLO_enugamma_pty_15_35	1.34×10^{2}	1.00	356 229	16 206 000	1.04×10^{7}	7.72×10^{1}
364523	Sherpa_222_NNPDF30NNLO_enugamma_pty_35_70	1.91×10^{1}	1.00	87 066	2 499 000	1.44×10^{6}	7.57×10^{1}
364524	Sherpa_222_NNPDF30NNLO_enugamma_pty_70_140	1.92	1.00	32 064	311 500	2.19×10^{5}	1.14×10^{2}
364525	Sherpa_222_NNPDF30NNLO_enugamma_pty_140_E_CMS	2.97×10^{-1}	1.00	19 211	121 500	9.97×10^{4}	3.35×10^{2}
364526	Sherpa_222_NNPDF30NNLO_munugamma_pty_7_15	1.99×10^{2}	1.00	291 176	10 792 000	6.63×10^{6}	3.32×10^{1}
364528	Sherpa_222_NNPDF30NNLO_munugamma_pty_35_70	1.91×10^{1}	1.00	102 516	2 499 000	1.45×10^{6}	7.57×10^{1}
364529	Sherpa_222_NNPDF30NNLO_munugamma_pty_70_140	1.92	1.00	32 811	312 500	2.20×10^{5}	1.15×10^{2}
364530	Sherpa_222_NNPDF30NNLO_munugamma_pty_140_E_CMS	2.98×10^{-1}	1.00	19 357	121 000	9.99×10^{4}	3.36×10^{2}
364531	Sherpa_222_NNPDF30NNLO_taunugamma_pty_7_15	1.99×10^{2}	1.00	25 206	15 623 000	9.59×10^{6}	4.81×10^{1}
364532	Sherpa_222_NNPDF30NNLO_taunugamma_pty_15_35	1.34×10^{2}	1.00	40 614	16 247 000	1.04×10^{7}	7.71×10^{1}
364533	Sherpa_222_NNPDF30NNLO_taunugamma_pty_35_70	1.91×10^{1}	1.00	14 594	2 499 000	1.45×10^{6}	7.60×10^{1}
364534	Sherpa_222_NNPDF30NNLO_taunugamma_pty_70_140	1.93	1.00	6511	309 500	2.17×10^{5}	1.13×10^{2}
364535	Sherpa_222_NNPDF30NNLO_taunugamma_pty_140_E_CMS	2.97×10^{-1}	1.00	5852	124 000	1.02×10^{5}	3.42×10^{2}
364500	Sherpa_222_NNPDF30NNLO_eegamma_pty_7_15	5.76×10^{1}	1.00	293 578	1 000 000	6.32×10^{5}	1.10×10^{1}
364501	Sherpa_222_NNPDF30NNLO_eegamma_pty_15_35	3.46×10^{1}	1.00	238 098	750 000	4.88×10^{5}	1.41×10^{1}
364502	Sherpa_222_NNPDF30NNLO_eegamma_pty_35_70	6.29	1.00	68 206	248 000	1.54×10^{5}	2.44×10^{1}
364503	Sherpa_222_NNPDF30NNLO_eegamma_pty_70_140	4.92×10^{-1}	1.00	58 313	123 000	8.59×10^{4}	1.75×10^{2}
364504	Sherpa_222_NNPDF30NNLO_eegamma_pty_140_E_CMS	6.30×10^{-2}	1.00	16 427	25 000	1.82×10^{4}	2.89×10^{2}
364505	Sherpa_222_NNPDF30NNLO_mumugamma_pty_7_15	5.77×10^{1}	1.00	375 634	1 000 000	6.32×10^{5}	1.09×10^{1}
364507	Sherpa_222_NNPDF30NNLO_mumugamma_pty_35_70	6.29	1.00	85 840	248 000	1.53×10^{5}	2.43×10^{1}
364508	Sherpa_222_NNPDF30NNLO_mumugamma_pty_70_140	4.94×10^{-1}	1.00	67 953	123 500	8.65×10^{4}	1.75×10^{2}
364509	Sherpa_222_NNPDF30NNLO_mumugamma_pty_140_E_CMS	6.31×10^{-2}	1.00	17 732	25 000	1.87×10^{4}	2.97×10^{2}
364510	Sherpa_222_NNPDF30NNLO_tautaugamma_pty_7_15	5.76×10^{1}	1.00	51 392	5 000 000	3.16×10^{6}	5.49×10^{1}
364511	Sherpa_222_NNPDF30NNLO_tautaugamma_pty_15_35	3.46×10^{1}	1.00	49 535	3 745 000	2.43×10^{6}	7.02×10^{1}
364512	Sherpa_222_NNPDF30NNLO_tautaugamma_pty_35_70	6.30	1.00	22 052	1 246 000	7.69×10^{5}	1.22×10^{2}
364513	Sherpa_222_NNPDF30NNLO_tautaugamma_pty_70_140	4.95×10^{-1}	1.00	5755	124 000	8.65×10^{4}	1.75×10^{2}
364514	Sherpa_222_NNPDF30NNLO_tautaugamma_pty_140_E_CMS	6.34×10^{-2}	1.00	10 227	100 000	7.42×10^{4}	1.17×10^{3}

Table 55: MC16 sub-campaign: MC16d. Summary of $V\gamma^{*}$ samples used in the analysis.

DSID	Name	$\sigma \cdot \epsilon$ [pb]	k-factor	N_{gen}	N_{all}	Σw	$\mathcal{L}$ [fb^{-1}]
361601	PowhegPy8EG_CT10nloME_AZNLOCTEQ6L1_WZlvll_mll4	4.51	1.00	2 561 281	10 010 000	4.51×10^{7}	9.99×10^{3}
363489	Sherpa_221_NNPDF30NNLO_WlvZqq	1.14×10^{1}	1.00	300 113	7 180 000	1.12×10^{6}	9.85×10^{1}
363358	Sherpa_221_NNPDF30NNLO_WqqZll	3.43	1.00	2 260 678	5 400 000	2.54×10^{5}	7.41×10^{1}

Table 56: MC16 sub-campaign: MC16a. Summary of WZ samples used in the analysis.

DSID	Name	$\sigma \cdot \epsilon$ [pb]	k-factor	N_{gen}	N_{all}	Σw	$\mathcal{L}$ [fb^{-1}]
361601	PowhegPy8EG_CT10nloME_AZNLOCTEQ6L1_WZlvll_mll4	4.51	1.00	2 462 000	10 001 000	4.51×10^{7}	1.00×10^{4}
363489	Sherpa_221_NNPDF30NNLO_WlvZqq	1.14×10^{1}	1.00	1 273 531	28 725 000	4.50×10^{6}	3.94×10^{2}
363358	Sherpa_221_NNPDF30NNLO_WqqZll	3.43	1.00	8 896 601	21 532 000	1.01×10^{6}	2.95×10^{2}

Table 57: MC16 sub-campaign: MC16d. Summary of WZ samples used in the analysis.

DSID	Name	$\sigma \cdot \epsilon$ [pb]	k-factor	N_{gen}	N_{all}	Σw	$\mathcal{L}$ [fb^{-1}]
361600	PowhegPy8EG_CT10nloME_AZNLOCTEQ6L1_WWlvlv	1.06×10^{1}	1.00	4 564 118	14 770 000	1.57×10^{8}	1.48×10^{4}
363360	Sherpa_221_NNPDF30NNLO_WplvWmqq	2.47×10^{1}	1.00	202 725	7 188 000	1.08×10^{6}	4.37×10^{1}
363359	Sherpa_221_NNPDF30NNLO_WpqqWmlv	2.47×10^{1}	1.00	202 423	7 194 000	1.08×10^{6}	4.36×10^{1}

Table 58: MC16 sub-campaign: MC16a. Summary of WW samples used in the analysis.

DSID	Name	$\sigma \cdot \epsilon$ [pb]	k-factor	N_{gen}	N_{all}	Σw	$\mathcal{L}$ [fb^{-1}]
361600	PowhegPy8EG_CT10nloME_AZNLOCTEQ6L1_WWlvlv	1.06×10^{1}	1.00	4 534 217	14 956 000	1.59×10^{8}	1.50×10^{4}
363360	Sherpa_221_NNPDF30NNLO_WplvWmqq	2.47×10^{1}	1.00	450 380	14 372 000	2.15×10^{6}	8.70×10^{1}
363359	Sherpa_221_NNPDF30NNLO_WpqqWmlv	2.47×10^{1}	1.00	909 512	28 711 000	4.29×10^{6}	1.74×10^{2}

Table 59: MC16 sub-campaign: MC16d. Summary of WW samples used in the analysis.

DSID	**Name**	$\sigma \cdot \epsilon$ [pb]	k-factor	N_{gen}	N_{all}	$\Sigma\, w$	$\mathcal{L}$ [fb^{-1}]
361603	PowhegPy8EG_CT10nloME_AZNLOCTEQ6L1_ZZllll_mll4	1.27	1.00	1 228 693	3 980 000	5.04×10^{6}	3.98×10^{3}
363356	Sherpa_221_NNPDF30NNLO_ZqqZll	2.21	1.00	2 352 582	5 400 000	3.49×10^{6}	1.58×10^{3}
361604	PowhegPy8EG_CT10nloME_AZNLOCTEQ6L1_ZZvvll_mll4	9.23×10^{-1}	1.00	280 660	995 000	9.13×10^{5}	9.90×10^{2}

Table 60: MC16 sub-campaign: MC16a. Summary of ZZ samples used in the analysis.

DSID	**Name**	$\sigma \cdot \epsilon$ [pb]	k-factor	N_{gen}	N_{all}	$\Sigma\, w$	$\mathcal{L}$ [fb^{-1}]
361603	PowhegPy8EG_CT10nloME_AZNLOCTEQ6L1_ZZllll_mll4	1.27	1.00	1 168 539	3 950 000	5.01×10^{6}	3.95×10^{3}
363356	Sherpa_221_NNPDF30NNLO_ZqqZll	2.21	1.00	2 325 180	5 400 000	3.49×10^{6}	1.58×10^{3}
361604	PowhegPy8EG_CT10nloME_AZNLOCTEQ6L1_ZZvvll_mll4	9.23×10^{-1}	1.00	274 889	995 000	9.19×10^{5}	9.96×10^{2}

Table 61: MC16 sub-campaign: MC16d. Summary of ZZ samples used in the analysis.

DSID	**Name**	$\sigma \cdot \epsilon$ [pb]	k-factor	N_{gen}	N_{all}	Σw	$\mathcal{L}$ [fb^{-1}]
407315	Sherpa_221_NNPDF30NNLO_2l4v_EW6	5.81×10^{-3}	1.00	21 748	60 000	6.01×10^{4}	1.03×10^{4}
407314	Sherpa_221_NNPDF30NNLO_3l3v_EW6	1.58×10^{-2}	1.00	29 559	50 000	5.03×10^{4}	3.19×10^{3}
407313	Sherpa_221_NNPDF30NNLO_4l2v_EW6	4.37×10^{-3}	1.00	34 096	50 000	5.01×10^{4}	1.15×10^{4}
407312	Sherpa_221_NNPDF30NNLO_5l1v_EW6	5.69×10^{-4}	1.00	35 019	50 000	5.02×10^{4}	8.83×10^{4}
407311	Sherpa_221_NNPDF30NNLO_6l0v_EW6	1.03×10^{-4}	1.00	33 025	50 000	5.00×10^{4}	4.88×10^{5}
341422	PowhegPythia8EvtGen_CT10_AZNLO_WmH125J_MINLO_eveWWlvqq	5.51×10^{-3}	1.00	20 907	50 000	2.93×10^{3}	5.31×10^{2}
341424	PowhegPythia8EvtGen_CT10_AZNLO_WmH125J_MINLO_muvmuWWlvqq	5.53×10^{-3}	1.00	21 966	50 000	2.94×10^{3}	5.32×10^{2}
341425	PowhegPythia8EvtGen_CT10_AZNLO_WmH125J_MINLO_qqWWlvlv	8.17×10^{-3}	1.00	15 165	50 000	1.80×10^{4}	2.21×10^{3}
341426	PowhegPythia8EvtGen_CT10_AZNLO_WmH125J_MINLO_qqWWlvqq	3.38×10^{-2}	1.00	1963	50 000	1.80×10^{4}	5.33×10^{2}
341428	PowhegPythia8EvtGen_CT10_AZNLO_WmH125J_MINLO_tauvtauWWlvqq	4.97×10^{-3}	1.00	6134	50 000	2.63×10^{3}	5.30×10^{2}
341430	PowhegPythia8EvtGen_CT10_AZNLO_WpH125J_MINLO_eveWWlvqq	7.97×10^{-3}	1.00	21 458	50 000	4.25×10^{3}	5.34×10^{2}
341432	PowhegPythia8EvtGen_CT10_AZNLO_WpH125J_MINLO_muvmuWWlvqq	8.00×10^{-3}	1.00	22 378	50 000	4.25×10^{3}	5.31×10^{2}
341433	PowhegPythia8EvtGen_CT10_AZNLO_WpH125J_MINLO_qqWWlvlv	1.20×10^{-2}	1.00	13 846	50 000	2.65×10^{4}	2.21×10^{3}
341434	PowhegPythia8EvtGen_CT10_AZNLO_WpH125J_MINLO_qqWWlvqq	5.01×10^{-2}	1.00	1829	50 000	2.66×10^{4}	5.31×10^{2}
341436	PowhegPythia8EvtGen_CT10_AZNLO_WpH125J_MINLO_tauvtauWWlvqq	8.02×10^{-3}	1.00	4555	40 000	3.31×10^{3}	4.13×10^{2}
341450	PowhegPythia8EvtGen_CT10_AZNLO_ZH125J_MINLO_eeWWlvqq_VpT	1.10×10^{-2}	1.00	75 232	90 000	2.18×10^{3}	1.98×10^{2}
341452	PowhegPythia8EvtGen_CT10_AZNLO_ZH125J_MINLO_mumuWWlvqq_VpT	1.16×10^{-2}	1.00	84 098	99 000	2.53×10^{3}	2.18×10^{2}
341454	PowhegPythia8EvtGen_CT10_AZNLO_ZH125J_MINLO_tautauWWlvqq_VpT	1.11×10^{-2}	1.00	27 692	100 000	2.47×10^{3}	2.22×10^{2}
341458	PowhegPythia8EvtGen_CT10_AZNLO_ZH125J_MINLO_vmuvmuWWlvqq_VpT	2.18×10^{-2}	1.00	1607	50 000	2.40×10^{3}	1.10×10^{2}

Table 62: MC16 sub-campaign: MC16a. Summary of triboson samples used in the analysis.

DSID	**Name**	$\sigma \cdot \epsilon$ [pb]	k-factor	N_{gen}	N_{all}	$\Sigma\, w$	$\mathcal{L}$ [fb^{-1}]
364244	Sherpa_222_NNPDF30NNLO_WWZ_2l4v_EW6	3.55×10^{-3}	1.00	23 667	60 000	1.38×10^{3}	3.90×10^{2}
364249	Sherpa_222_NNPDF30NNLO_ZZZ_2l4v_EW6	1.71×10^{-4}	1.00	16 823	35 000	1.08×10^{3}	6.30×10^{3}
364242	Sherpa_222_NNPDF30NNLO_WWW_3l3v_EW6	7.20×10^{-3}	1.00	38 753	60 000	2.28×10^{3}	3.17×10^{2}
364246	Sherpa_222_NNPDF30NNLO_WZZ_3l3v_EW6	7.44×10^{-4}	1.00	32 111	50 000	1.63×10^{3}	2.19×10^{3}
364243	Sherpa_222_NNPDF30NNLO_WWZ_4l2v_EW6	1.80×10^{-3}	1.00	47 716	60 000	6.97×10^{2}	3.88×10^{2}
364248	Sherpa_222_NNPDF30NNLO_ZZZ_4l2v_EW6	8.65×10^{-5}	1.00	27 686	35 000	1.08×10^{3}	1.25×10^{4}
364245	Sherpa_222_NNPDF30NNLO_WZZ_5l1v_EW6	1.88×10^{-4}	1.00	43 718	50 000	1.83×10^{2}	9.75×10^{2}
364247	Sherpa_222_NNPDF30NNLO_ZZZ_6l0v_EW6	1.45×10^{-5}	1.00	32 226	35 000	4.00×10^{1}	2.82×10^{3}
341422	PowhegPythia8EvtGen_CT10_AZNLO_eveWWlvqq	5.51×10^{-3}	1.00	20 392	50 000	2.92×10^{3}	5.30×10^{2}
341424	PowhegPythia8EvtGen_CT10_AZNLO_muvmuWWlvqq	5.53×10^{-3}	1.00	21 666	50 000	2.93×10^{3}	5.29×10^{2}
341425	PowhegPythia8EvtGen_CT10_AZNLO_qqWWlvlv	8.17×10^{-3}	1.00	14 515	50 000	1.83×10^{4}	2.24×10^{3}
341426	PowhegPythia8EvtGen_CT10_AZNLO_qqWWlvqq	3.38×10^{-2}	1.00	1948	50 000	1.82×10^{4}	5.40×10^{2}
341428	PowhegPythia8EvtGen_CT10_AZNLO_tauvtauWWlvqq	4.97×10^{-3}	1.00	4991	45 000	2.64×10^{3}	5.31×10^{2}
341430	PowhegPythia8EvtGen_CT10_AZNLO_eveWWlvqq	7.97×10^{-3}	1.00	20 351	50 000	4.74×10^{3}	5.94×10^{2}
341432	PowhegPythia8EvtGen_CT10_AZNLO_muvmuWWlvqq	8.00×10^{-3}	1.00	21 862	50 000	4.73×10^{3}	5.91×10^{2}
341433	PowhegPythia8EvtGen_CT10_AZNLO_qqWWlvlv	1.20×10^{-2}	1.00	12 326	45 000	2.66×10^{4}	2.21×10^{3}
341434	PowhegPythia8EvtGen_CT10_AZNLO_qqWWlvqq	5.01×10^{-2}	1.00	1864	50 000	2.96×10^{4}	5.90×10^{2}
341436	PowhegPythia8EvtGen_CT10_AZNLO_tauvtauWWlvqq	8.02×10^{-3}	1.00	5558	50 000	4.73×10^{3}	5.90×10^{2}
341450	PowhegPythia8EvtGen_CT10_AZNLO_eeWWlvqq_VpT	1.10×10^{-2}	1.00	71 933	87 000	2.23×10^{3}	2.02×10^{2}
341452	PowhegPythia8EvtGen_CT10_AZNLO_mumuWWlvqq_VpT	1.16×10^{-2}	1.00	84 233	99 500	2.69×10^{3}	2.32×10^{2}
341454	PowhegPythia8EvtGen_CT10_AZNLO_tautauWWlvqq_VpT	1.11×10^{-2}	1.00	26 795	100 000	2.58×10^{3}	2.32×10^{2}
341456	PowhegPythia8EvtGen_CT10_AZNLO_veveWWlvqq_VpT	2.18×10^{-2}	1.00	1565	45 000	2.28×10^{3}	1.05×10^{2}
341458	PowhegPythia8EvtGen_CT10_AZNLO_vmuvmuWWlvqq_VpT	2.18×10^{-2}	1.00	1661	50 000	2.55×10^{3}	1.17×10^{2}

Table 63: MC16 sub-campaign: MC16d. Summary of triboson samples used in the analysis.

B Binning distribution for the optimization study

In order to determine which variable is best to use as binning for the significance in the optimization study, the significances obtained when applying four different binning variables are compared. The four variables are: $M(\mu_1, \mu_2, e)$; $M(\mu_2, e)$; $p_T(\mu_2)$; and $p_{\mathrm{T}}(e)$. Each variable is related to the topology of the signal: $M(\mu_1, \mu_2, e)$ is correlated with the W-boson mass; and both $M(\mu_2, e)$, $p_T(\mu_2)$ and $p_{\mathrm{T}}(e)$ are correlated with M_N. The four distributions are shown in Fig. 47 for the signal region applying the preselection requirements, see Tab. 8.

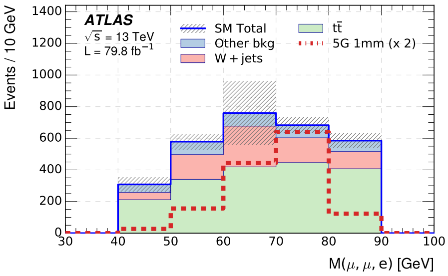
The significance is determined using Eq. 7.2. The MC statistical uncertainty is determined using a sum over the MC weights. When comparing the distributions, a bin number of $N_{\mathrm{bins}} = 5$ and a systematic uncertainty of $p_{\mathrm{syst}} = 0.5$ is applied. The ranges used for the variables are:

$M(\mu_1, \mu_2, e)$:	Interval from 40-90 GeV	See distribution in Fig. 47a and 47b
$M(\mu_2, e)$:	Interval from 0-50 GeV	See distribution in Fig. 47c and 47d
$p_T(\mu_2)$:	Interval from 0-50 GeV	See distribution in Fig. 47e and 47f
$p_{\mathrm{T}}(e)$:	Interval from 0-50 GeV	See distribution in Fig. 47g and 47h

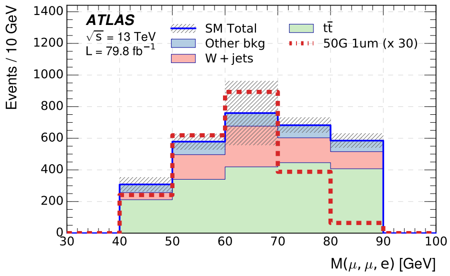
Figure 48-50 show the optimization plots for the variables $p_T(\mu_1)$, $p_T(\mu_2)$ and $p_T(e)$ respectively using each of the four binning distributions for the signal samples with $M_N = 5$ GeV and $M_N = 50$ GeV. The estimated significance is clearly affected by which variable is used to bin the distribution. In bins where no signal events are present, the significance contribution is zero and thereby not affected by a decrease or increase in the number of background events.

When applying the variables $p_T(\mu_2)$ and $p_{\mathrm{T}}(e)$ to bin the significance, the significance is clearly affected by the mass of the HNL. When M_N is small, $p_T(\mu_2)$ and $p_T(e)$ are in general smaller than for the signal samples with higher M_N. The background contributing to the signal region is peaked at low p_T for the subleading muon and electron, so when determining the significance using binning in $p_T(\mu_2)$ and $p_T(e)$, a relatively low significance is estimated for signal samples with low M_N and high significance for the signal samples with high M_N. The shape of the $M(\mu_2, e)$ distribution is also highly dependent on the value of M_N for the signal sample since in general $M(\mu_2, e) < M_N$. This means that for $M_N = 5$ GeV, it is only the first bin of the distribution for which $s > 0$ when applying $N_{\mathrm{bins}} = 5$ and thereby effectively a one-bin significance. For an increasing size of M_N for the signal sample, the effective number of bins contributing is increasing too. For the distribution of $M(\mu_1, \mu_2, e)$, the shape of the signal samples is not very dependent on M_N. When M_N is low, a smaller amount of energy is in general missing due to the SM neutrino of the final state, so the variable $M(\mu_1, \mu_2, e)$ tends to be peaking closer to the mass of the W-boson at 80 GeV.

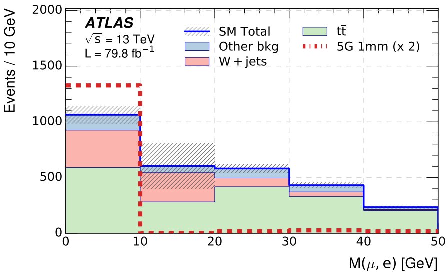
In order to avoid a large influence of the mass M_N for the different signal samples on the estimation of the significance, the distribution $M(\mu_1, \mu_2, e)$ is used to bin the significance in the optimization study.



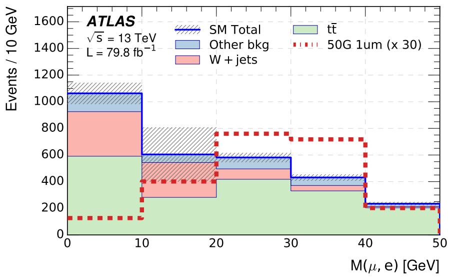
(a) Signal region, $M(\mu_1, \mu_2, e)$ distribution. Signal with $M_N = 5$ GeV.



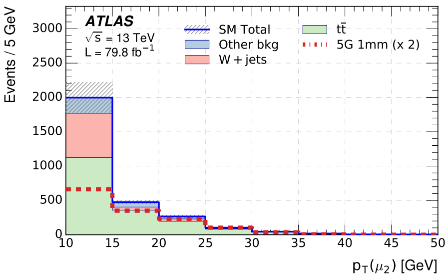
(b) Signal region, $M(\mu_1, \mu_2, e)$ distribution. Signal with $M_N = 50$ GeV.



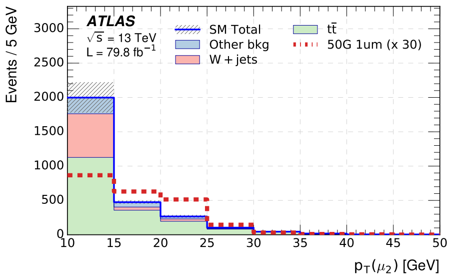
(c) Signal region, $M(\mu_2, e)$ distribution. Signal with $M_N = 5$ GeV.



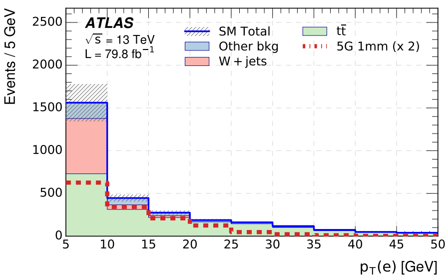
(d) Signal region, $M(\mu_2, e)$ distribution. Signal with $M_N = 50$ GeV.



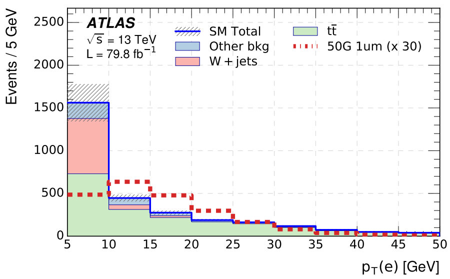
(e) Signal region, $p_T(\mu_2)$ distribution. Signal with $M_N = 5$ GeV.



(f) Signal region, $p_T(\mu_2)$ distribution. Signal with $M_N = 50$ GeV.

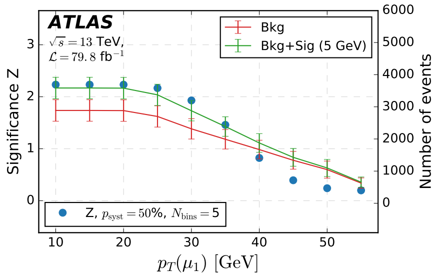


(g) Signal region, $p_T(e)$ distribution. Signal with $M_N = 5$ GeV.

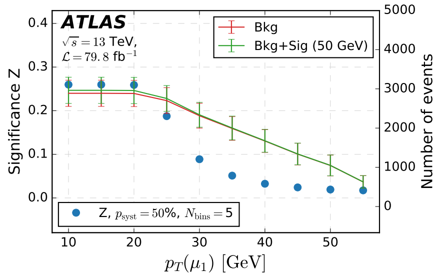


(h) Signal region, $p_T(e)$ distribution. Signal with $M_N = 50$ GeV.

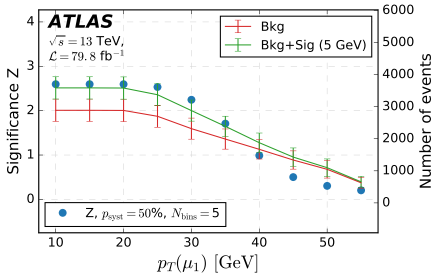
Figure 47: Plots of the four distributions which are possibly being applied for binning when determining the significance. The distributions are showed in the signal region with $M_N = 5$ and 50 GeV. Preselection requirements are applied, see Tabl. 8.



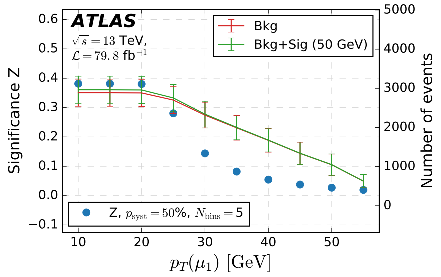
(a) $M_N = 5$ GeV. Binning variable $M(\mu_1, \mu_2, e)$.



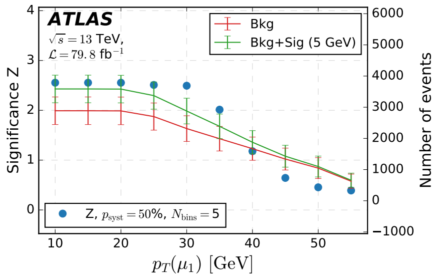
(b) $M_N = 50$ GeV. Binning variable $M(\mu_1, \mu_2, e)$.



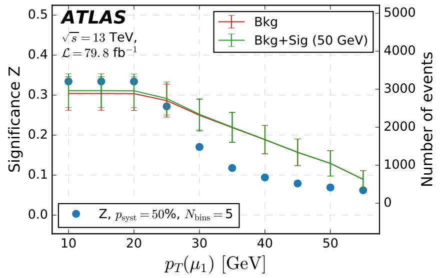
(c) $M_N = 5$ GeV. Binning variable $M(\mu_2, e)$.



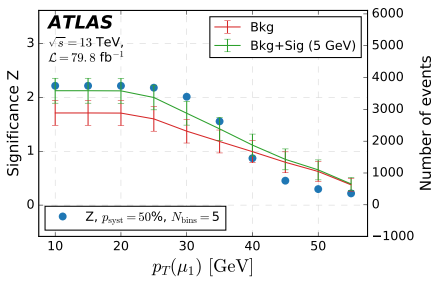
(d) $M_N = 50$ GeV. Binning variable $M(\mu_2, e)$.



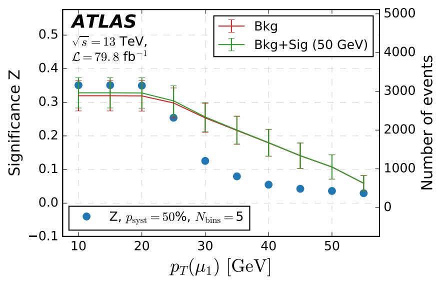
(e) $M_N = 5$ GeV. Binning variable $p_T(\mu_2)$.



(f) $M_N = 50$ GeV. Binning variable $p_T(\mu_2)$.

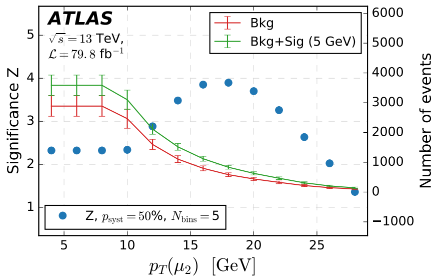


(g) $M_N = 5$ GeV. Binning variable $p_T(e)$.

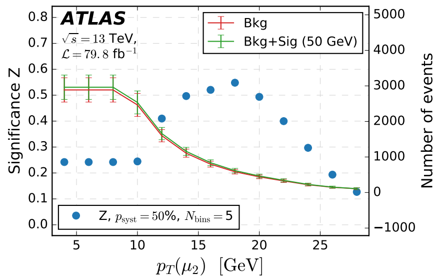


(h) $M_N = 50$ GeV. Binning variable $p_T(e)$.

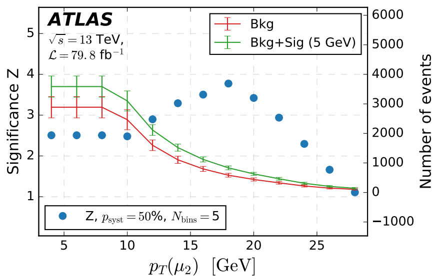
Figure 48: Optimization study for the variable $p_T(\mu_1)$. The plots show the significance and number of background and signal events as a function of the cut on $p_T(\mu_1)$ for which the x-axis value x_0 represents the requirement $p_T(\mu_1) \geq x_0$ GeV.



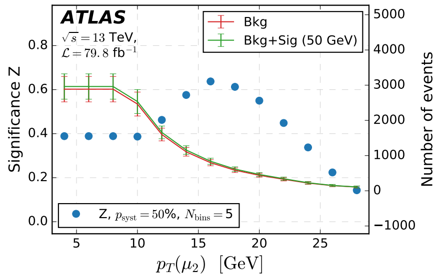
(a) $M_N = 5$ GeV. Binning variable $M(\mu_1, \mu_2, e)$.



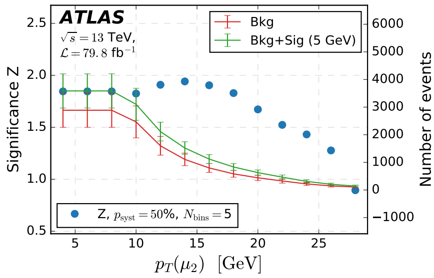
(b) $M_N = 50$ GeV. Binning variable $M(\mu_1, \mu_2, e)$.



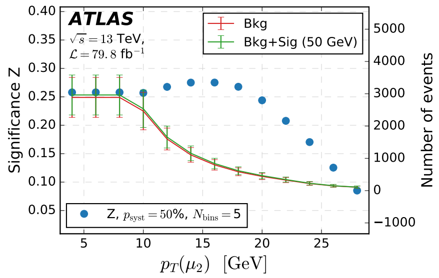
(c) $M_N = 5$ GeV. Binning variable $M(\mu_2, e)$.



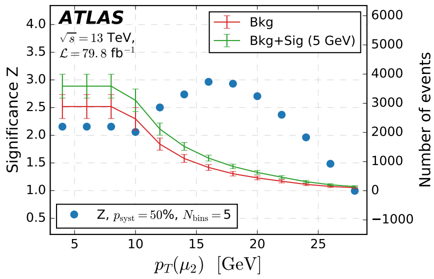
(d) $M_N = 50$ GeV. Binning variable $M(\mu_2, e)$.



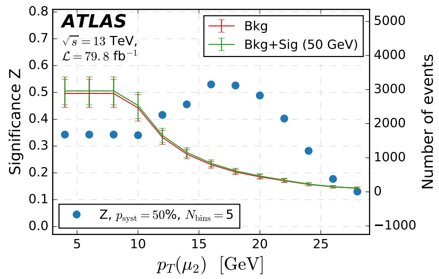
(e) $M_N = 5$ GeV. Binning variable $p_T(\mu_2)$.



(f) $M_N = 50$ GeV. Binning variable $p_T(\mu_2)$.

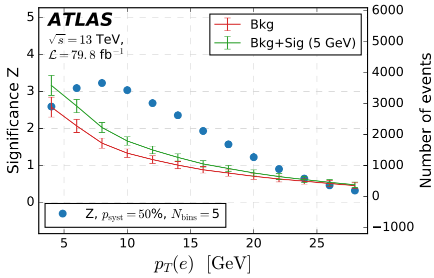


(g) $M_N = 5$ GeV. Binning variable $p_T(e)$.

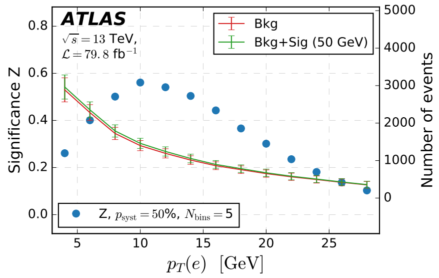


(h) $M_N = 50$ GeV. Binning variable $p_T(e)$.

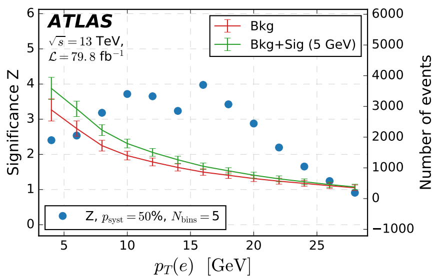
Figure 49: Optimization study for the variable $p_T(\mu_2)$. The plots show the significance and number of background and signal events as a function of the cut on $p_T(\mu_2)$ for which the x-axis value x_0 represents the requirement $p_T(\mu_2) \geq x_0$ GeV.



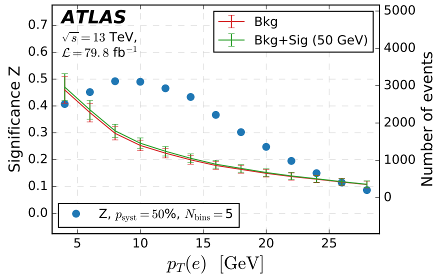
(a) $M_N = 5$ GeV. Binning variable $M(\mu_1, \mu_2, e)$.



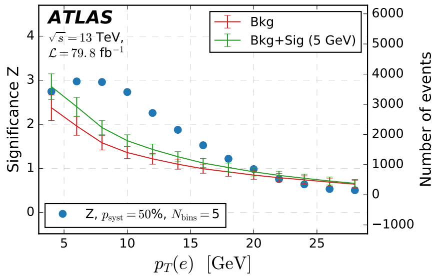
(b) $M_N = 50$ GeV. Binning variable $M(\mu_1, \mu_2, e)$.



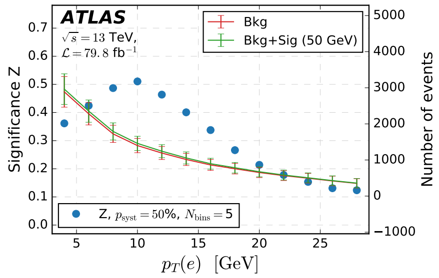
(c) $M_N = 5$ GeV. Binning variable $M(\mu_2, e)$.



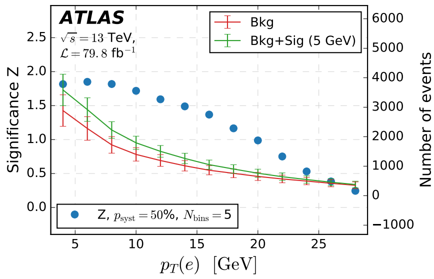
(d) $M_N = 50$ GeV. Binning variable $M(\mu_2, e)$.



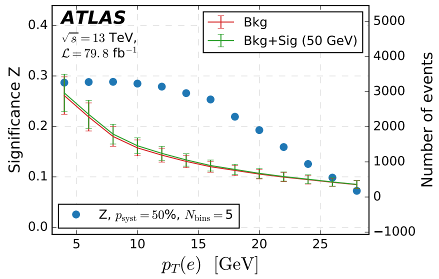
(e) $M_N = 5$ GeV. Binning variable $p_T(\mu_2)$.



(f) $M_N = 50$ GeV. Binning variable $p_T(\mu_2)$.



(g) $M_N = 5$ GeV. Binning variable $p_T(e)$.



(h) $M_N = 50$ GeV. Binning variable $p_T(e)$.

Figure 50: Optimization study for the variable $p_T(e)$. The plots show the significance and number of background and signal events as a function of the cut on $p_T(e)$ for which the x-axis value x_0 represents the requirement $p_T(e) \geq x_0$ GeV.

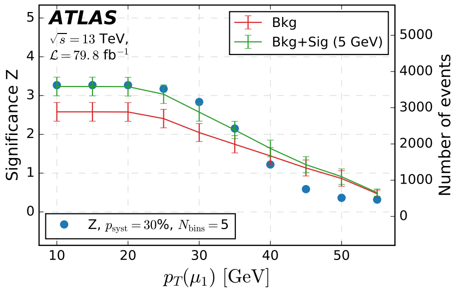
C Optimization study

The plots applied in the optimization study of the signal region requirements are shown in this section. Plots are included for three of the five signal samples with $M_N = 5$, 20 and 50 GeV. In Sec. 7.3, the signal region optimization study is performed by examining the effect of a range of different cuts on the significance and number of background and signal events. Only a few selected plots were included directly in the optimization study in Sec. 7.3, so in this section a more complete list of the analysed plots is shown.

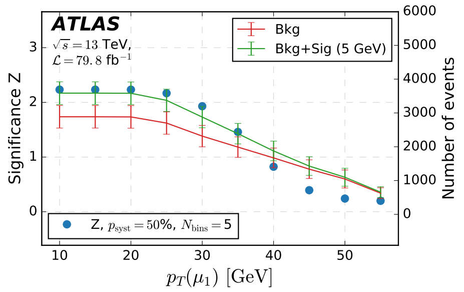
For each variable considered in the optimization study, plots are included for the three signal samples $M_N = 5$, 20 and 50 GeV. For each signal sample, the significance scan is done for the four combinations of: $N_{\text{bins}} = 5/10$ and $p_{\text{syst}} = 0.3/0.5$. Table 64 shows the list of variables considered in the optimization study, the signal region requirements chosen for each variable, and the figures provided for the three signal samples: $M_N = 5$, 20 and 50 GeV.

Variable	**SR requirement**	**Figures of optimization study**
$p_T(\mu_1)$	$p_T(\mu_1) \geq 23$ GeV	Fig. 51 (5 GeV), 52 (20 GeV), 53 (50 GeV)
$p_T(e)$	$p_T(e) \geq 10$ GeV	Fig. 54 (5 GeV), 55 (20 GeV), 56 (50 GeV)
$p_T(\mu_2)$	$p_T(\mu_2) \geq 10$ GeV	Fig. 57 (5 GeV), 58 (20 GeV), 59 (50 GeV)
$N_{\text{b-jets}}$	b-jet veto: $N_{\text{b-jets}} = 0$	Fig. 60 (5 GeV), 61 (20 GeV), 62 (50 GeV)
Isolation for μ_1	`FixedCutTightTrackOnly`	Fig. 63 (5 GeV), 64 (20 GeV), 65 (50 GeV)
Isolation for e	`Gradient`	Fig. 66 (5 GeV), 67 (20 GeV), 68 (50 GeV)
Isolation for μ_2	`FixedCutLoose`	Fig. 69 (5 GeV), 70 (20 GeV), 71 (50 GeV)
E_T^{miss}	$E_T^{\text{miss}} < 60$ GeV	Fig. 72 (5 GeV), 73 (20 GeV), 74 (50 GeV)
Identification for μ_1	`Tight`	Fig. 75 (5 GeV), 76 (20 GeV), 77 (50 GeV)
$\Delta R(\mu_2, e)$	$\Delta R(\mu_2, e) > 0.2$	Fig. 78 (5 GeV), 79 (20 GeV), 80 (50 GeV)
$\|d_0\|/\sigma(d_0)(e)$	$\|d_0\|/\sigma(d_0)(e) < 10$	Fig. 81 (5 GeV), 82 (20 GeV), 83 (50 GeV)
$\|d_0\|/\sigma(d_0)(\mu_2)$	$\|d_0\|/\sigma(d_0)(\mu_2) < 10$	Fig. 84 (5 GeV), 85 (20 GeV), 86 (50 GeV)
$\|\Delta z_0 \sin(\theta)\|(e)$	$\|\Delta z_0 \sin(\theta)\|(e) < 0.5$ mm	Fig. 87 (5 GeV), 88 (20 GeV), 89 (50 GeV)
$\|\Delta z_0 \sin(\theta)\|(\mu_2)$	$\|\Delta z_0 \sin(\theta)\|(\mu_2) < 0.5$ mm	Fig. 90 (5 GeV), 91 (20 GeV), 92 (50 GeV)

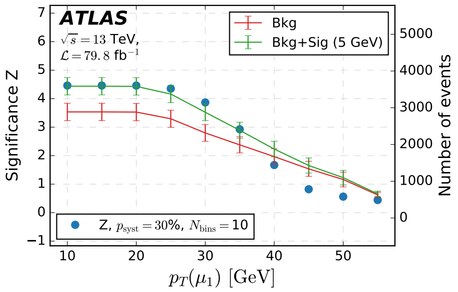
Table 64: The list of variables examined in the optimization study of the signal region requirements. Prior to the optimization study, the preselection requirements listed in Tab. 8 are applied. The signal region requirements are listed in chronological order, in the sense that prior to the optimization study of a variable both the preselection requirements and all the signal region requirements listed before the study are applied.



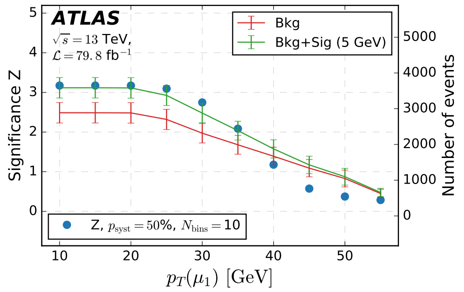
(a) $N_{\text{bins}} = 5$, $\text{p}^{\text{syst}} = 30\%$



(b) $N_{\text{bins}} = 5$, $\text{p}^{\text{syst}} = 50\%$

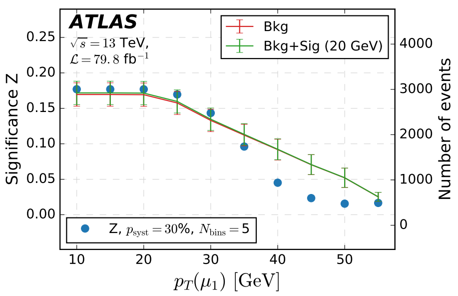


(c) $N_{\text{bins}} = 10$, $\text{p}^{\text{syst}} = 30\%$

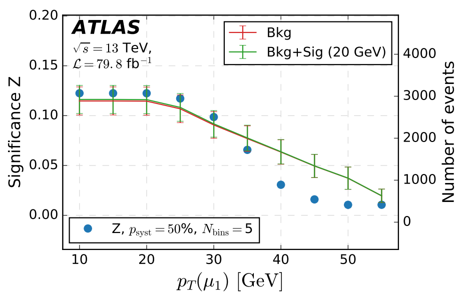


(d) $N_{\text{bins}} = 10$, $\text{p}^{\text{syst}} = 50\%$

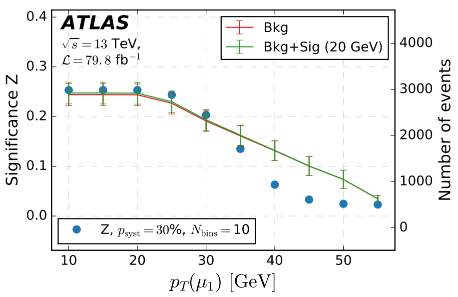
Figure 51: Signal sample with $M_N = 5$ GeV. Optimization study for the variable $p_T(\mu_1)$. The plots show the significance and number of background and signal events as a function of the cut on $p_T(\mu_1)$ for which the x-axis value x_0 represents the requirement $p_T(\mu_1) \geq x_0$ GeV.



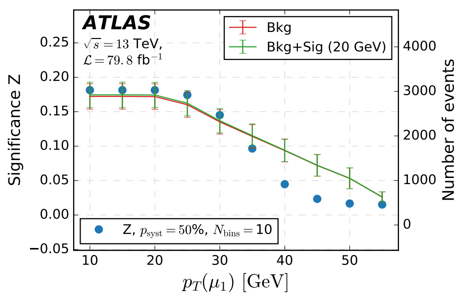
(a) $N_{\text{bins}} = 5$, $\text{p}^{\text{syst}} = 30\%$



(b) $N_{\text{bins}} = 5$, $\text{p}^{\text{syst}} = 50\%$

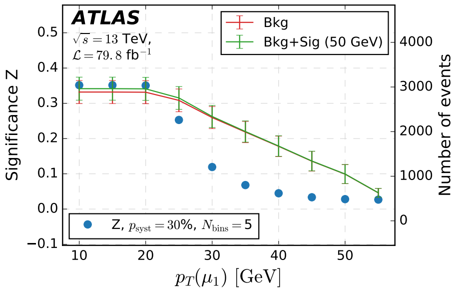


(c) $N_{\text{bins}} = 10$, $\text{p}^{\text{syst}} = 30\%$

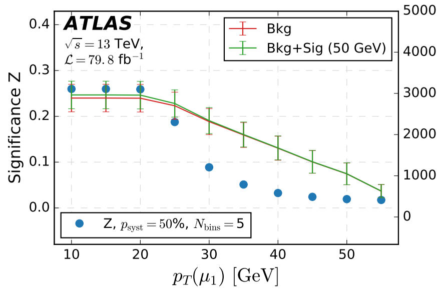


(d) $N_{\text{bins}} = 10$, $\text{p}^{\text{syst}} = 50\%$

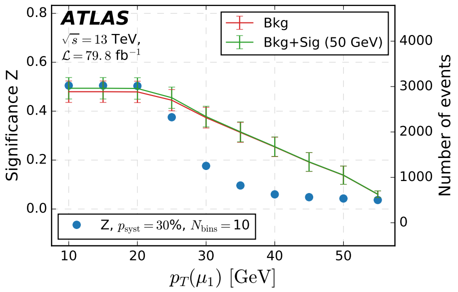
Figure 52: Signal sample with $M_N = 20$ GeV. Optimization study for the variable $p_T(\mu_1)$. The plots show the significance and number of background and signal events as a function of the cut on $p_T(\mu_1)$ for which the x-axis value x_0 represents the requirement $p_T(\mu_1) \geq x_0$ GeV.



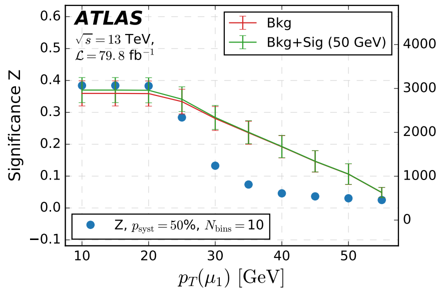
(a) $N_{\rm bins} = 5$, $\mathrm{p}^{\rm syst} = 30\%$



(b) $N_{\rm bins} = 5$, $\mathrm{p}^{\rm syst} = 50\%$

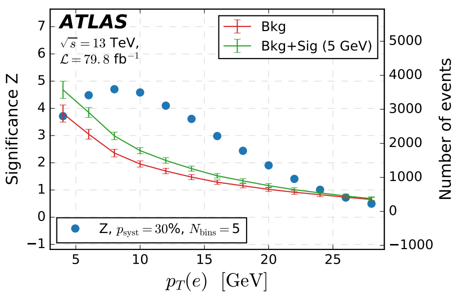


(c) $N_{\rm bins} = 10$, $\mathrm{p}^{\rm syst} = 30\%$

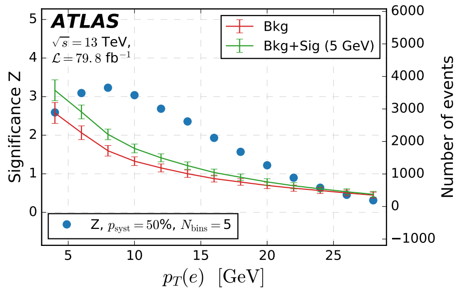


(d) $N_{\rm bins} = 10$, $\mathrm{p}^{\rm syst} = 50\%$

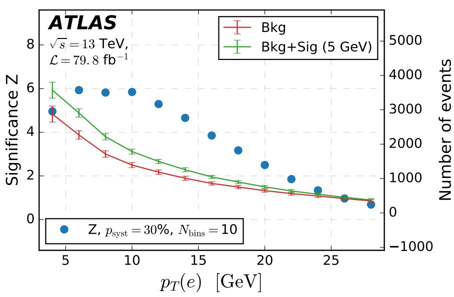
Figure 53: Signal sample with $M_N = 50$ GeV. Optimization study for the variable $p_T(\mu_1)$. The plots show the significance and number of background and signal events as a function of the cut on $p_T(\mu_1)$ for which the x-axis value x_0 represents the requirement $p_T(\mu_1) \geq x_0$ GeV.



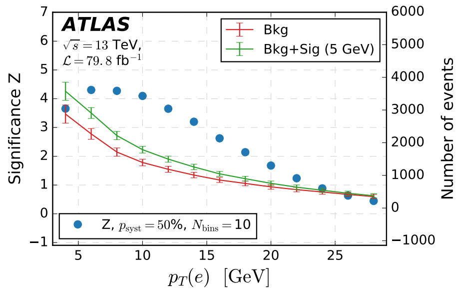
(a) $N_{\rm bins} = 5$, $\mathrm{p}^{\rm syst} = 30\%$



(b) $N_{\rm bins} = 5$, $\mathrm{p}^{\rm syst} = 50\%$

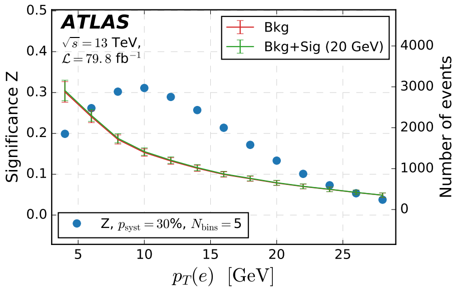


(c) $N_{\rm bins} = 10$, $\mathrm{p}^{\rm syst} = 30\%$

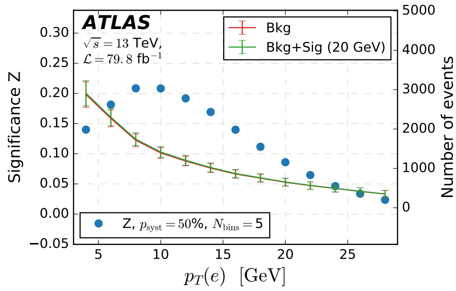


(d) $N_{\rm bins} = 10$, $\mathrm{p}^{\rm syst} = 50\%$

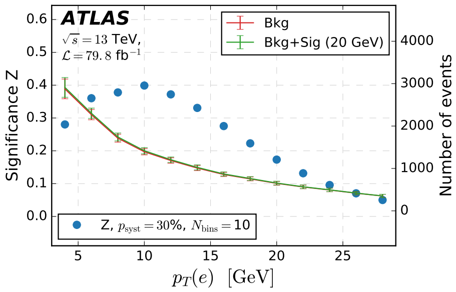
Figure 54: Signal sample with $M_N = 5$ GeV. Optimization study for the variable $p_T(e)$. The plots show the significance and number of background and signal events as a function of the cut on $p_T(e)$ for which the x-axis value x_0 represents the requirement $p_T(e) \geq x_0$ GeV.



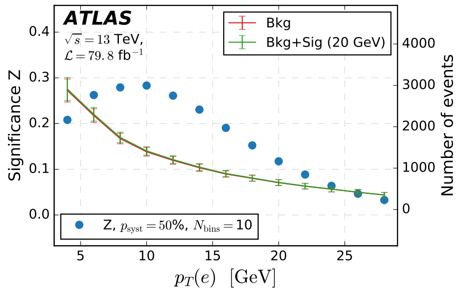
(a) $N_{\text{bins}} = 5$, $\text{p}^{\text{syst}} = 30\%$



(b) $N_{\text{bins}} = 5$, $\text{p}^{\text{syst}} = 50\%$

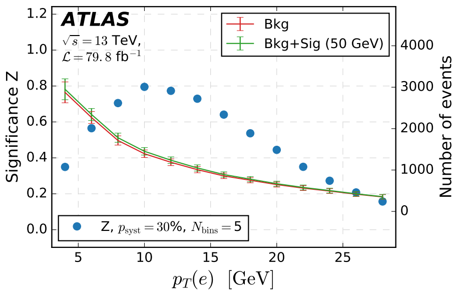


(c) $N_{\text{bins}} = 10$, $\text{p}^{\text{syst}} = 30\%$

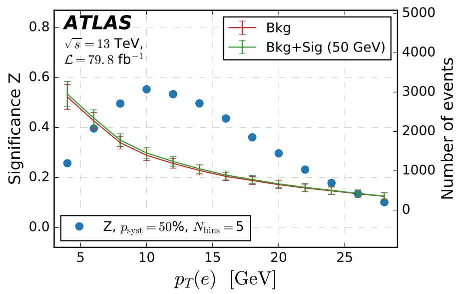


(d) $N_{\text{bins}} = 10$, $\text{p}^{\text{syst}} = 50\%$

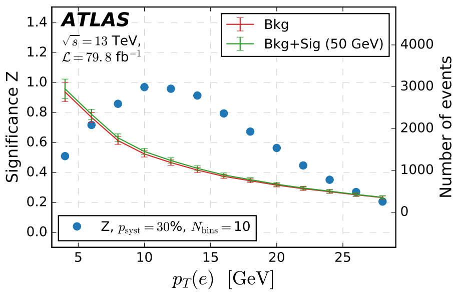
Figure 55: Signal sample with $M_N = 20$ GeV. Optimization study for the variable $p_T(e)$. The plots show the significance and number of background and signal events as a function of the cut on $p_T(e)$ for which the x-axis value x_0 represents the requirement $p_T(e) \geq x_0$ GeV.



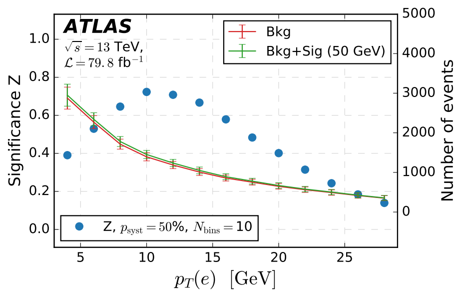
(a) $N_{\text{bins}} = 5$, $\text{p}^{\text{syst}} = 30\%$



(b) $N_{\text{bins}} = 5$, $\text{p}^{\text{syst}} = 50\%$

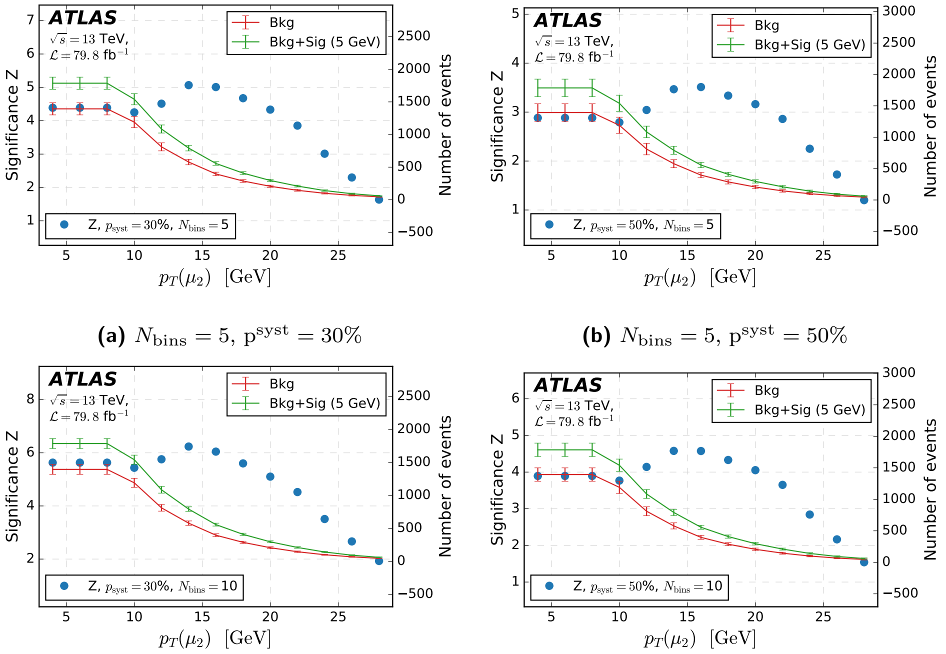


(c) $N_{\text{bins}} = 10$, $\text{p}^{\text{syst}} = 30\%$



(d) $N_{\text{bins}} = 10$, $\text{p}^{\text{syst}} = 50\%$

Figure 56: Signal sample with $M_N = 50$ GeV. Optimization study for the variable $p_T(e)$. The plots show the significance and number of background and signal events as a function of the cut on $p_T(e)$ for which the x-axis value x_0 represents the requirement $p_T(e) \geq x_0$ GeV.



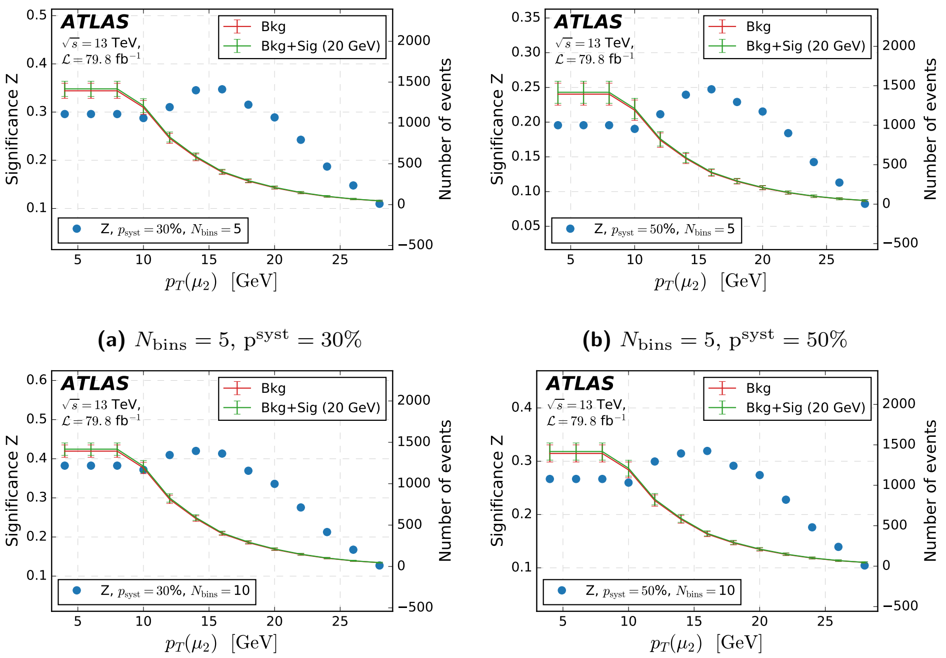
(a) $N_{\text{bins}} = 5$, $\text{p}^{\text{syst}} = 30\%$

(b) $N_{\text{bins}} = 5$, $\text{p}^{\text{syst}} = 50\%$

(c) $N_{\text{bins}} = 10$, $\text{p}^{\text{syst}} = 30\%$

(d) $N_{\text{bins}} = 10$, $\text{p}^{\text{syst}} = 50\%$

Figure 57: Signal sample with $M_N = 5$ GeV. Optimization study for the variable $p_T(\mu_2)$. The plots show the significance and number of background and signal events as a function of the cut on $p_T(\mu_2)$ for which the x-axis value x_0 represents the requirement $p_T(\mu_2) \geq x_0$ GeV.



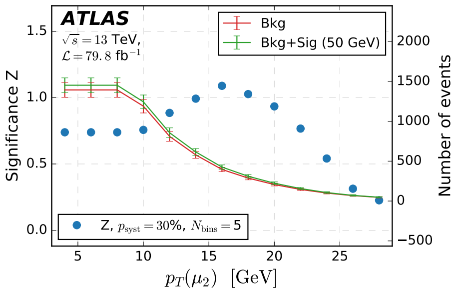
(a) $N_{\text{bins}} = 5$, $\text{p}^{\text{syst}} = 30\%$

(b) $N_{\text{bins}} = 5$, $\text{p}^{\text{syst}} = 50\%$

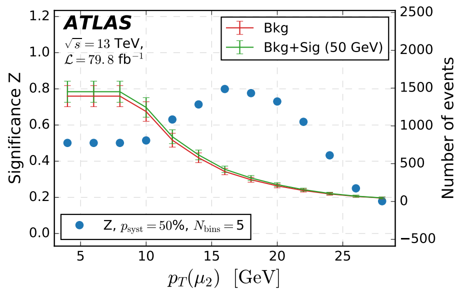
(c) $N_{\text{bins}} = 10$, $\text{p}^{\text{syst}} = 30\%$

(d) $N_{\text{bins}} = 10$, $\text{p}^{\text{syst}} = 50\%$

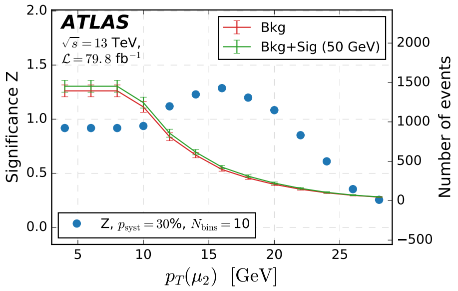
Figure 58: Signal sample with $M_N = 20$ GeV. Optimization study for the variable $p_T(\mu_2)$. The plots show the significance and number of background and signal events as a function of the cut on $p_T(\mu_2)$ for which the x-axis value x_0 represents the requirement $p_T(\mu_2) \geq x_0$ GeV.



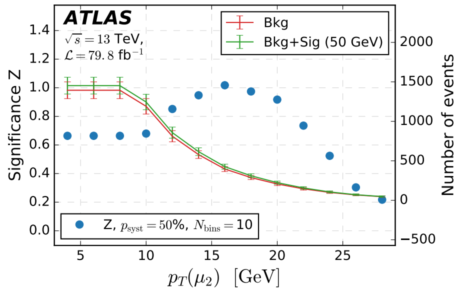
(a) $N_{\text{bins}} = 5$, $\text{p}^{\text{syst}} = 30\%$



(b) $N_{\text{bins}} = 5$, $\text{p}^{\text{syst}} = 50\%$

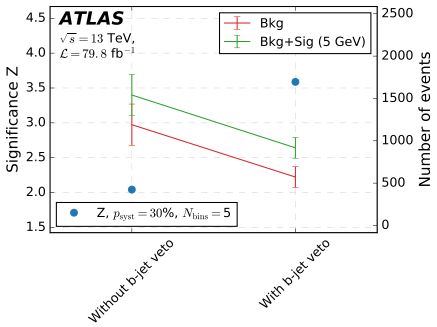


(c) $N_{\text{bins}} = 10$, $\text{p}^{\text{syst}} = 30\%$

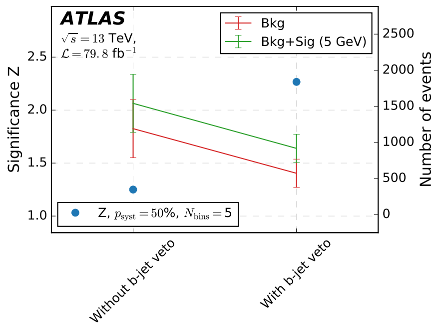


(d) $N_{\text{bins}} = 10$, $\text{p}^{\text{syst}} = 50\%$

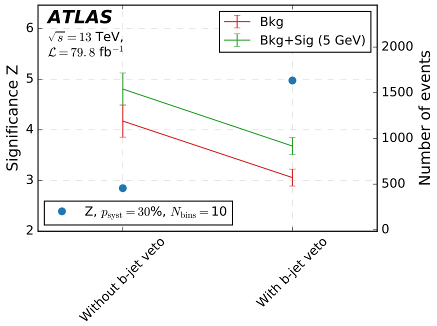
Figure 59: Signal sample with $M_N = 50$ GeV. Optimization study for the variable $p_T(\mu_2)$. The plots show the significance and number of background and signal events as a function of the cut on $p_T(\mu_2)$ for which the x-axis value x_0 represents the requirement $p_T(\mu_2) \geq x_0$ GeV.



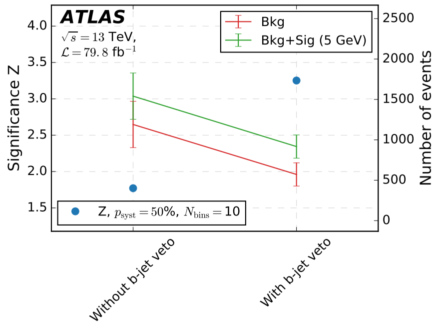
(a) $N_{\text{bins}} = 5$, $\text{p}^{\text{syst}} = 30\%$



(b) $N_{\text{bins}} = 5$, $\text{p}^{\text{syst}} = 50\%$

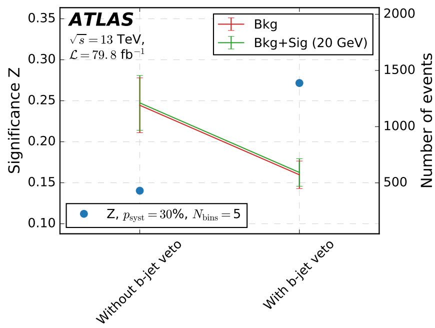


(c) $N_{\text{bins}} = 10$, $\text{p}^{\text{syst}} = 30\%$

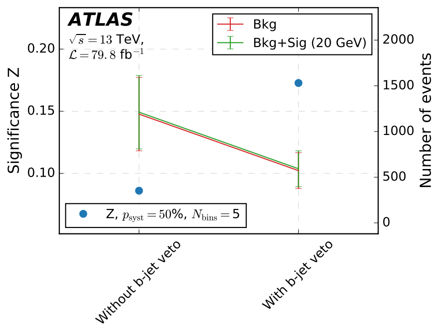


(d) $N_{\text{bins}} = 10$, $\text{p}^{\text{syst}} = 50\%$

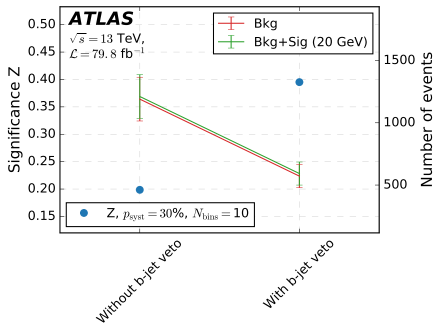
Figure 60: Signal sample with $M_N = 5$ GeV. Optimization study for the variable $N_{\text{b-jets}}$. The plots show the significance and number of background and signal events as a function of the cut on $N_{\text{b-jets}}$.



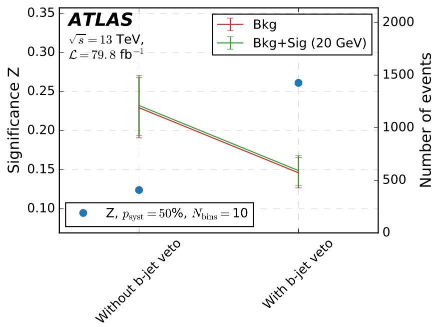
(a) $N_{\text{bins}} = 5$, $\text{p}^{\text{syst}} = 30\%$



(b) $N_{\text{bins}} = 5$, $\text{p}^{\text{syst}} = 50\%$

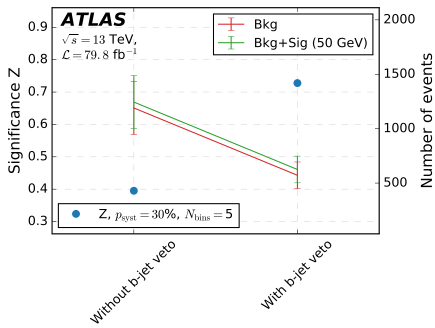


(c) $N_{\text{bins}} = 10$, $\text{p}^{\text{syst}} = 30\%$

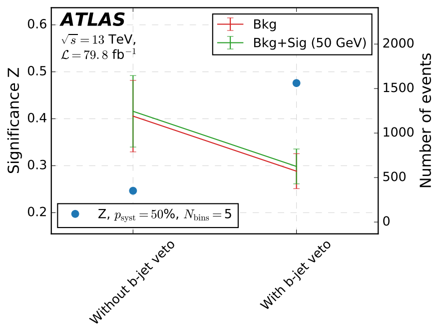


(d) $N_{\text{bins}} = 10$, $\text{p}^{\text{syst}} = 50\%$

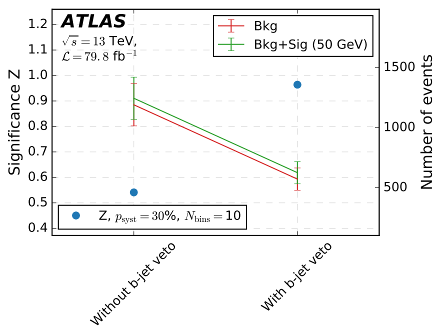
Figure 61: Signal sample with $M_N = 20$ GeV. Optimization study for the variable $N_{\text{b-jets}}$. The plots show the significance and number of background and signal events as a function of the cut on $N_{\text{b-jets}}$.



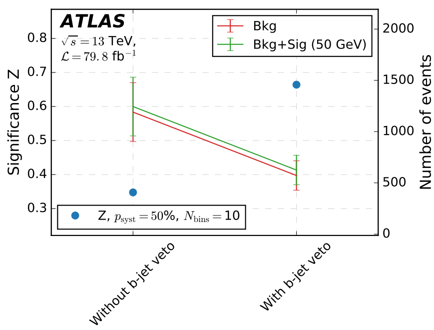
(a) $N_{\text{bins}} = 5$, $\text{p}^{\text{syst}} = 30\%$



(b) $N_{\text{bins}} = 5$, $\text{p}^{\text{syst}} = 50\%$

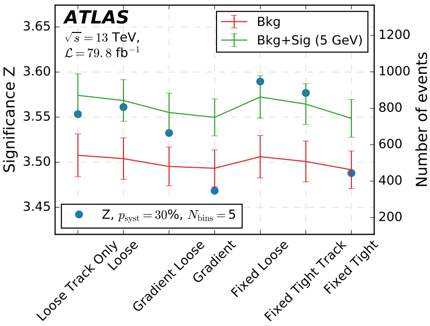


(c) $N_{\text{bins}} = 10$, $\text{p}^{\text{syst}} = 30\%$

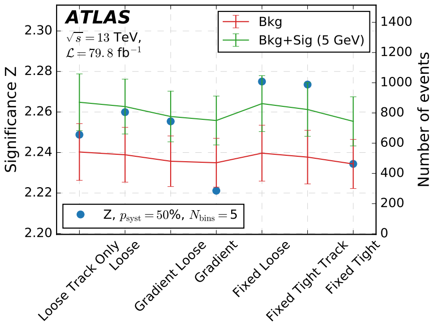


(d) $N_{\text{bins}} = 10$, $\text{p}^{\text{syst}} = 50\%$

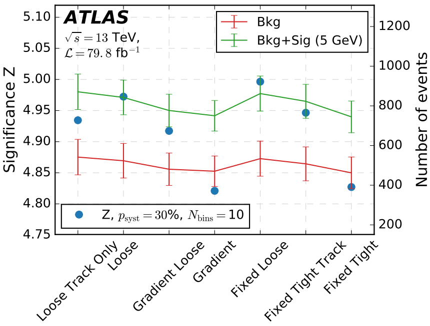
Figure 62: Signal sample with $M_N = 50$ GeV. Optimization study for the variable $N_{\text{b-jets}}$. The plots show the significance and number of background and signal events as a function of the cut on $N_{\text{b-jets}}$.



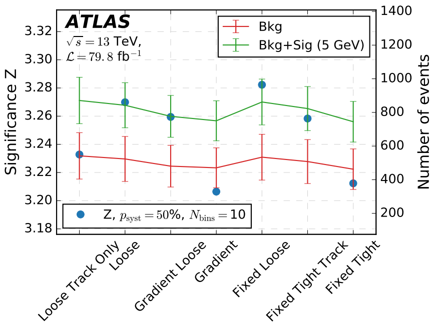
(a) $N_{\text{bins}} = 5$, $\text{p}^{\text{syst}} = 30\%$



(b) $N_{\text{bins}} = 5$, $\text{p}^{\text{syst}} = 50\%$

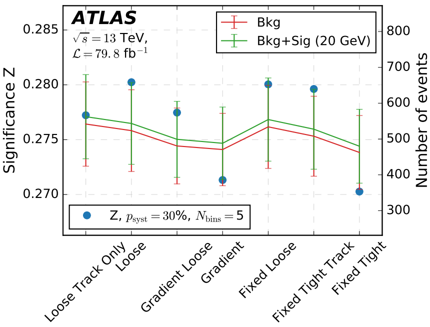


(c) $N_{\text{bins}} = 10$, $\text{p}^{\text{syst}} = 30\%$

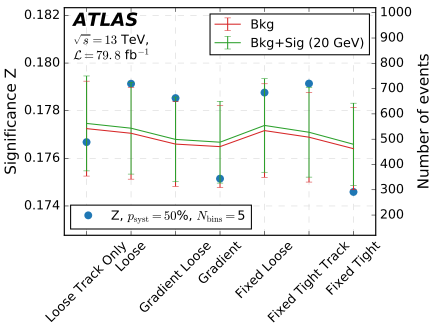


(d) $N_{\text{bins}} = 10$, $\text{p}^{\text{syst}} = 50\%$

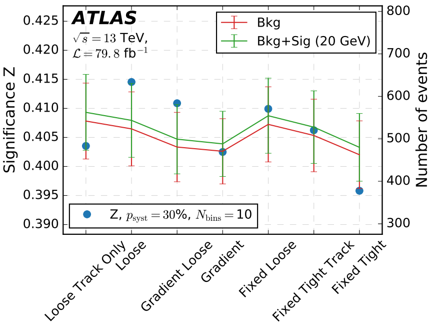
Figure 63: Signal sample with $M_N = 5$ GeV. Optimization study for the isolation selection of μ_1. The plots show the significance and number of background and signal events as a function of the isolation selection of μ_1.



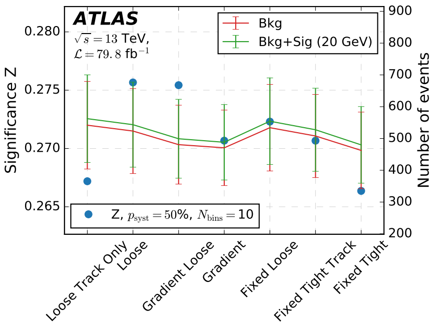
(a) $N_{\text{bins}} = 5$, $\text{p}^{\text{syst}} = 30\%$



(b) $N_{\text{bins}} = 5$, $\text{p}^{\text{syst}} = 50\%$

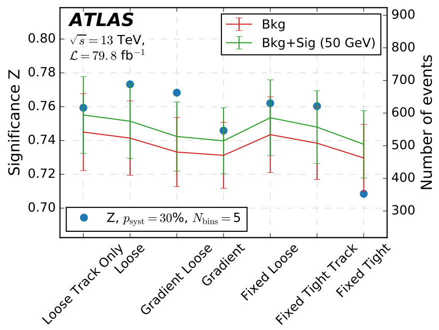


(c) $N_{\text{bins}} = 10$, $\text{p}^{\text{syst}} = 30\%$

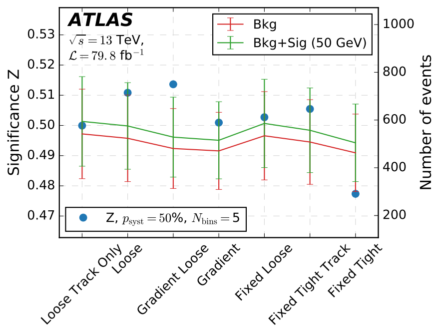


(d) $N_{\text{bins}} = 10$, $\text{p}^{\text{syst}} = 50\%$

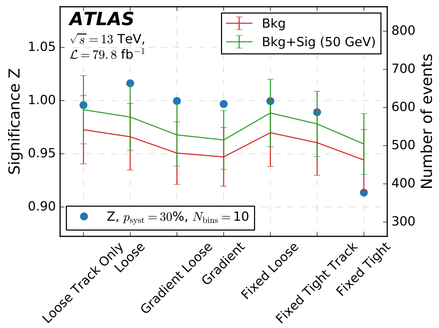
Figure 64: Signal sample with $M_N = 20$ GeV. Optimization study for the isolation selection of μ_1. The plots show the significance and number of background and signal events as a function of the isolation selection of μ_1.



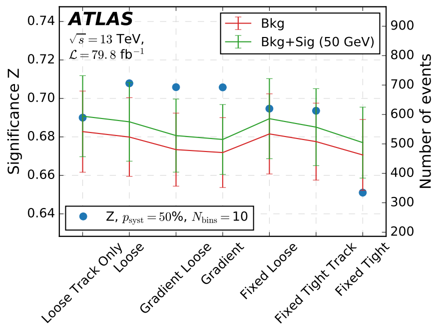
(a) $N_{\rm bins}=5$, $\mathrm{p}^{\rm syst}=30\%$



(b) $N_{\rm bins}=5$, $\mathrm{p}^{\rm syst}=50\%$

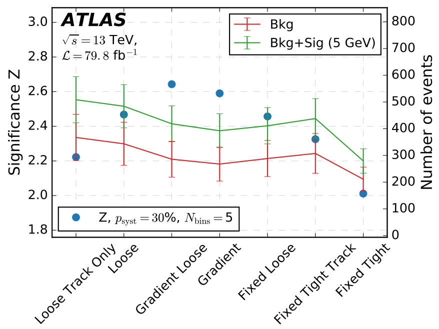


(c) $N_{\rm bins}=10$, $\mathrm{p}^{\rm syst}=30\%$

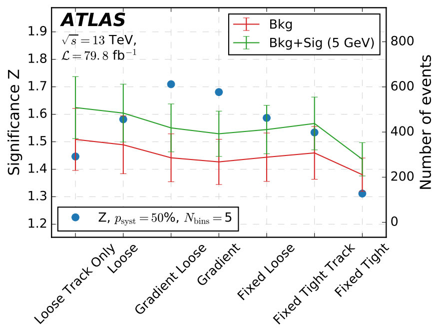


(d) $N_{\rm bins}=10$, $\mathrm{p}^{\rm syst}=50\%$

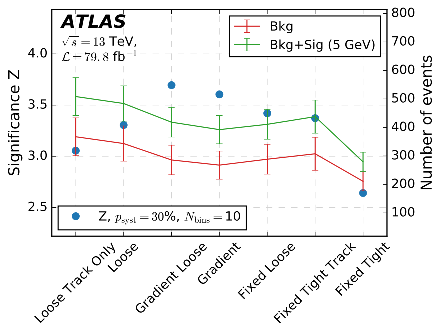
Figure 65: Signal sample with $M_N = 50$ GeV. Optimization study for the isolation selection of μ_1. The plots show the significance and number of background and signal events as a function of the the isolation selection of μ_1.



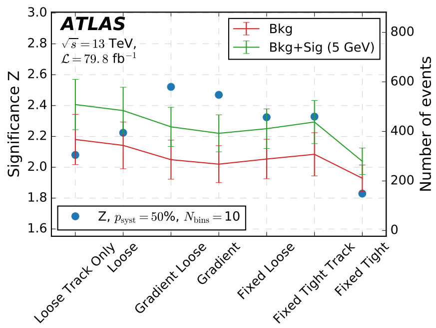
(a) $N_{\rm bins}=5$, $\mathrm{p}^{\rm syst}=30\%$



(b) $N_{\rm bins}=5$, $\mathrm{p}^{\rm syst}=50\%$

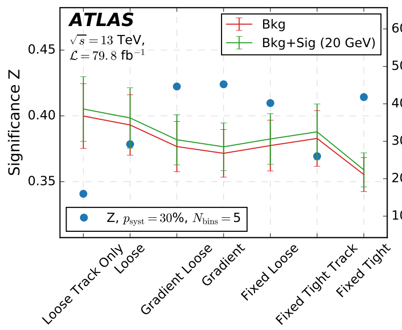


(c) $N_{\rm bins}=10$, $\mathrm{p}^{\rm syst}=30\%$

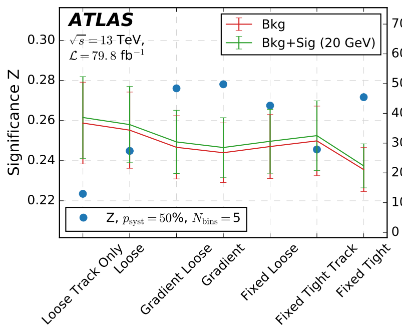


(d) $N_{\rm bins}=10$, $\mathrm{p}^{\rm syst}=50\%$

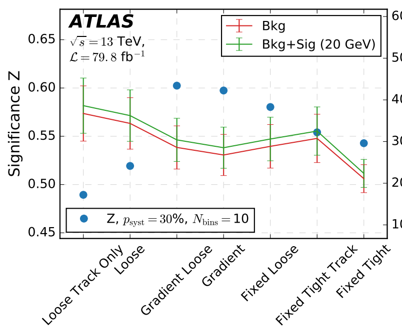
Figure 66: Signal sample with $M_N = 5$ GeV. Optimization study for the isolation selection of e. The plots show the significance and number of background and signal events as a function of the the isolation selection of e.



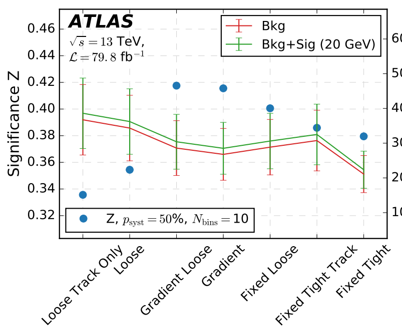
(a) $N_{\rm bins} = 5$, $\rm p^{syst} = 30\%$



(b) $N_{\rm bins} = 5$, $\rm p^{syst} = 50\%$

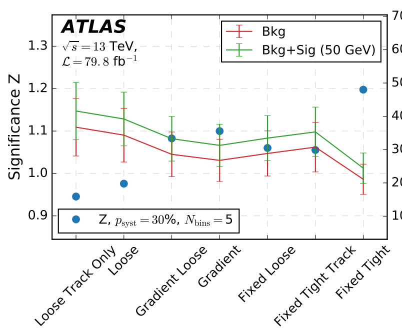


(c) $N_{\rm bins} = 10$, $\rm p^{syst} = 30\%$

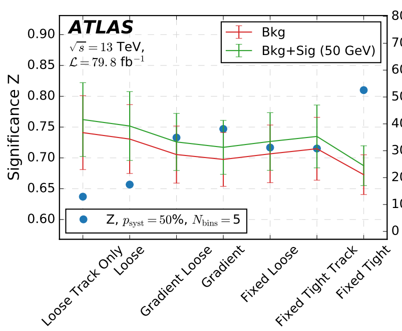


(d) $N_{\rm bins} = 10$, $\rm p^{syst} = 50\%$

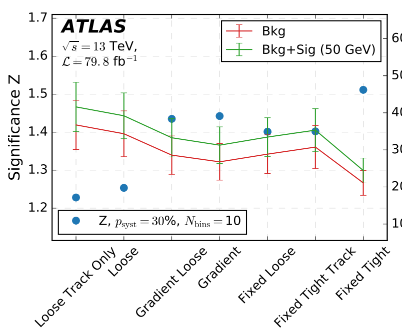
Figure 67: Signal sample with $M_N = 20$ GeV. Optimization study for the isolation selection of e. The plots show the significance and number of background and signal events as a function of the isolation selection of e.



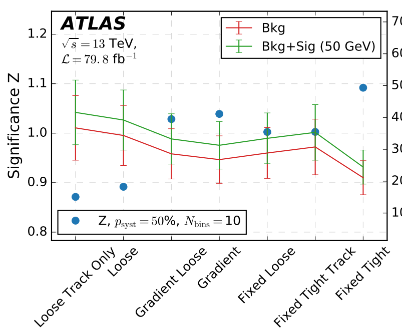
(a) $N_{\rm bins} = 5$, $\rm p^{syst} = 30\%$



(b) $N_{\rm bins} = 5$, $\rm p^{syst} = 50\%$

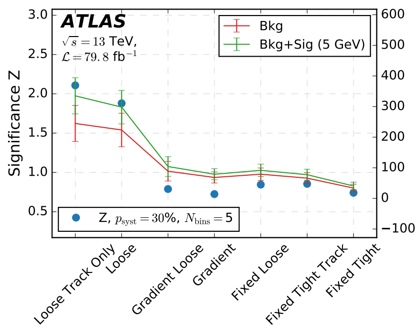


(c) $N_{\rm bins} = 10$, $\rm p^{syst} = 30\%$

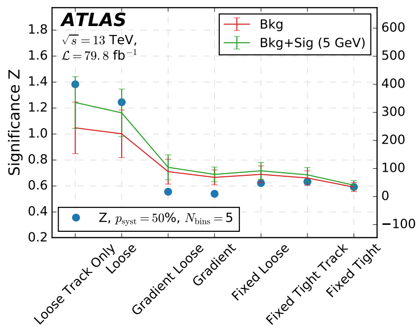


(d) $N_{\rm bins} = 10$, $\rm p^{syst} = 50\%$

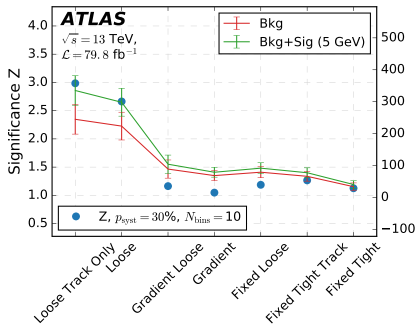
Figure 68: Signal sample with $M_N = 50$ GeV. Optimization study for the isolation selection of e. The plots show the significance and number of background and signal events as a function of the the isolation selection of e.



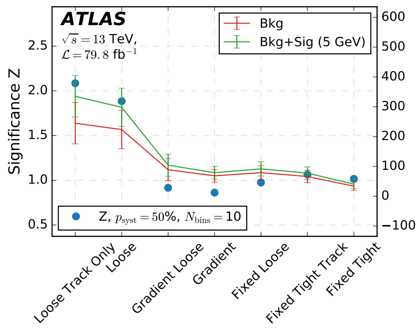
(a) $N_{\rm bins} = 5$, $\rm p^{syst} = 30\%$



(b) $N_{\rm bins} = 5$, $\rm p^{syst} = 50\%$

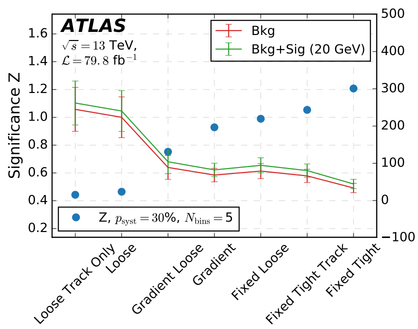


(c) $N_{\rm bins} = 10$, $\rm p^{syst} = 30\%$

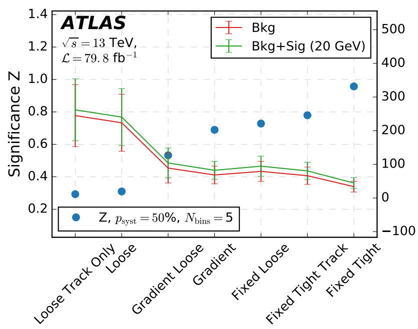


(d) $N_{\rm bins} = 10$, $\rm p^{syst} = 50\%$

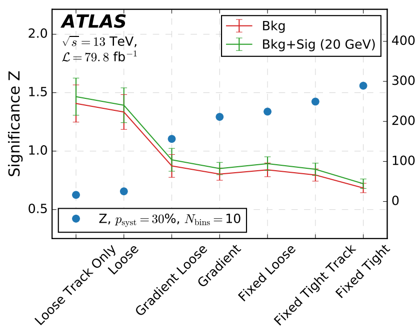
Figure 69: Signal sample with $M_N = 5$ GeV. Optimization study for the isolation selection of μ_2. The plots show the significance and number of background and signal events as a function of the isolation selection of μ_2.



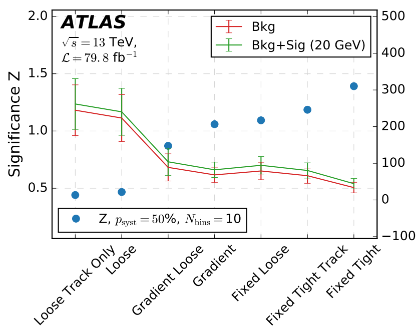
(a) $N_{\rm bins} = 5$, $\rm p^{syst} = 30\%$



(b) $N_{\rm bins} = 5$, $\rm p^{syst} = 50\%$

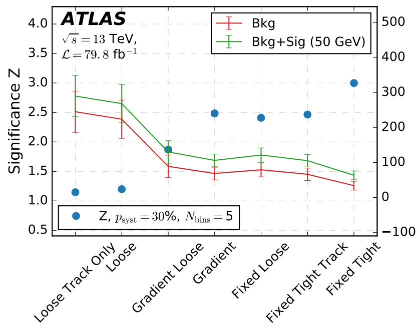


(c) $N_{\rm bins} = 10$, $\rm p^{syst} = 30\%$

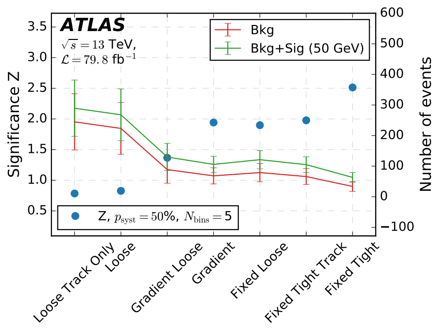


(d) $N_{\rm bins} = 10$, $\rm p^{syst} = 50\%$

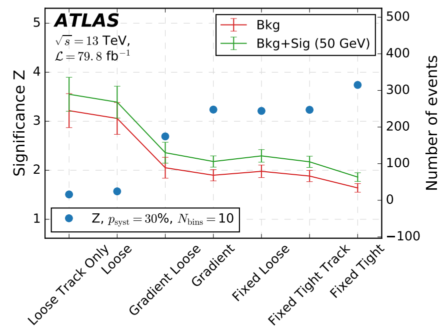
Figure 70: Signal sample with $M_N = 20$ GeV. Optimization study for the isolation selection of μ_2. The plots show the significance and number of background and signal events as a function of the isolation selection of μ_2.



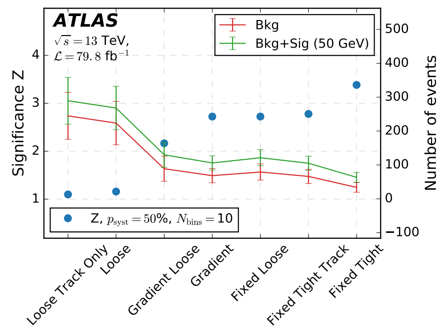
(a) $N_{\rm bins}=5$, $\mathrm{p}^{\rm syst}=30\%$



(b) $N_{\rm bins}=5$, $\mathrm{p}^{\rm syst}=50\%$

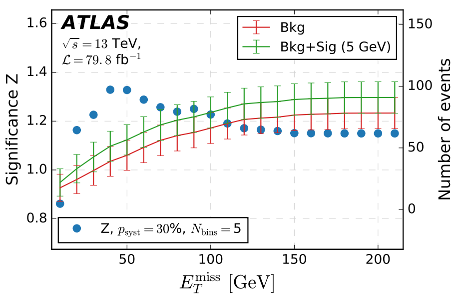


(c) $N_{\rm bins}=10$, $\mathrm{p}^{\rm syst}=30\%$

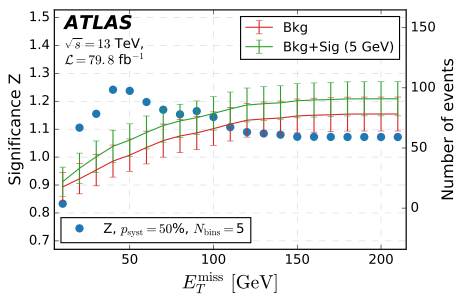


(d) $N_{\rm bins}=10$, $\mathrm{p}^{\rm syst}=50\%$

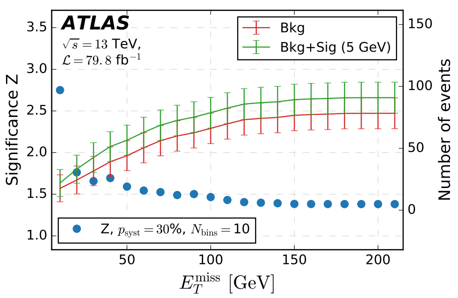
Figure 71: Signal sample with $M_N = 50$ GeV. Optimization study for the isolation selection of μ_2. The plots show the significance and number of background and signal events as a function of the the isolation selection of μ_2.



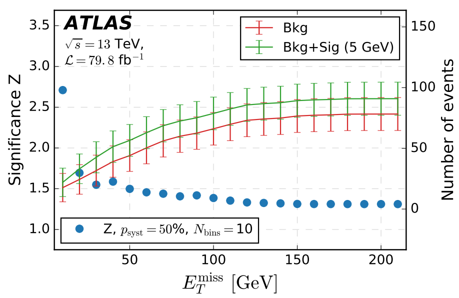
(a) $N_{\rm bins}=5$, $\mathrm{p}^{\rm syst}=30\%$



(b) $N_{\rm bins}=5$, $\mathrm{p}^{\rm syst}=50\%$

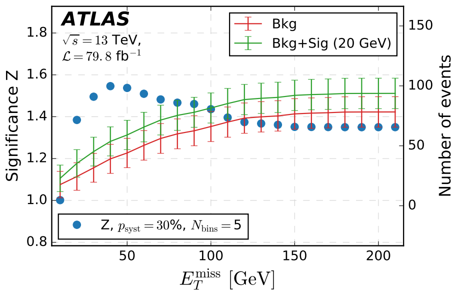


(c) $N_{\rm bins}=10$, $\mathrm{p}^{\rm syst}=30\%$

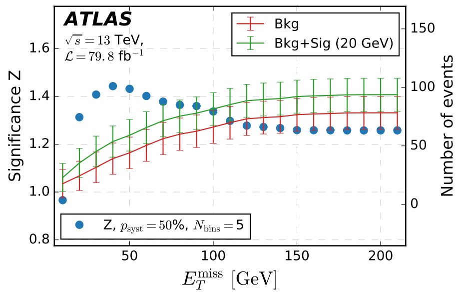


(d) $N_{\rm bins}=10$, $\mathrm{p}^{\rm syst}=50\%$

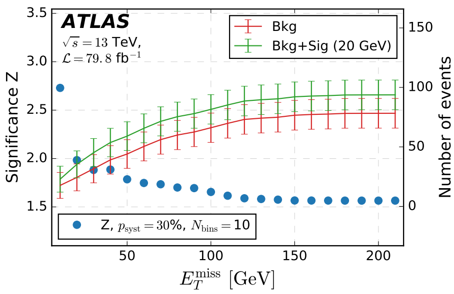
Figure 72: Signal sample with $M_N = 5$ GeV. Optimization study for the variable $E_T^{\rm miss}$. The plots show the significance and number of background and signal events as a function of the cut on $E_T^{\rm miss}$ for which the x-axis value x_0 represents the requirement $E_T^{\rm miss} \le x_0$ GeV.



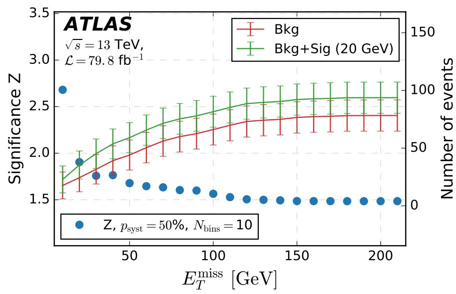
(a) $N_{\text{bins}} = 5$, $\text{p}^{\text{syst}} = 30\%$



(b) $N_{\text{bins}} = 5$, $\text{p}^{\text{syst}} = 50\%$

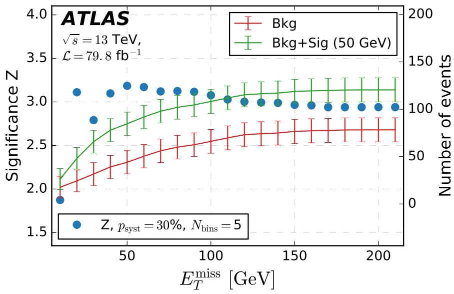


(c) $N_{\text{bins}} = 10$, $\text{p}^{\text{syst}} = 30\%$

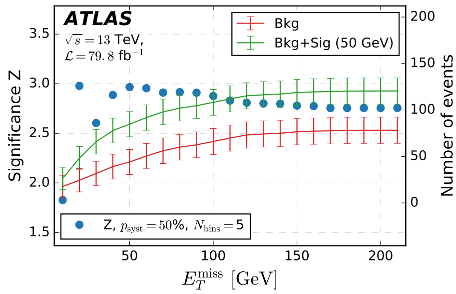


(d) $N_{\text{bins}} = 10$, $\text{p}^{\text{syst}} = 50\%$

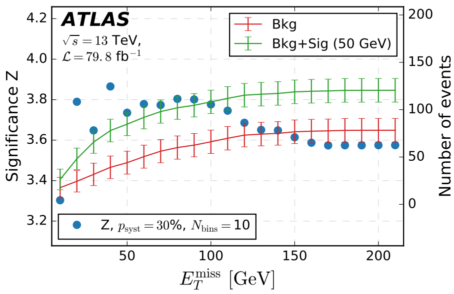
Figure 73: Signal sample with $M_N = 20$ GeV. Optimization study for the variable E_T^{miss}. The plots show the significance and number of background and signal events as a function of the cut on E_T^{miss} for which the x-axis value x_0 represents the requirement $E_T^{\text{miss}} \leq x_0$ GeV.



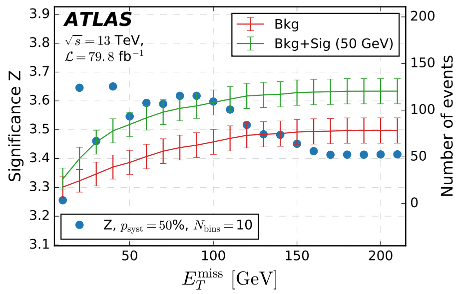
(a) $N_{\text{bins}} = 5$, $\text{p}^{\text{syst}} = 30\%$



(b) $N_{\text{bins}} = 5$, $\text{p}^{\text{syst}} = 50\%$

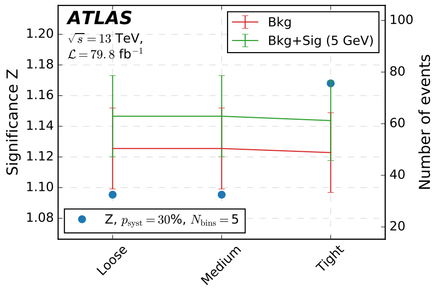


(c) $N_{\text{bins}} = 10$, $\text{p}^{\text{syst}} = 30\%$

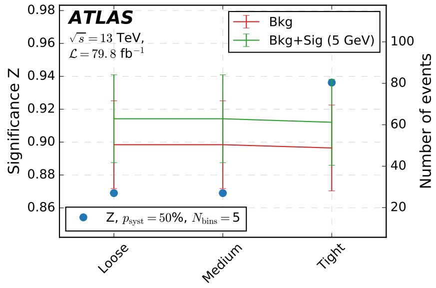


(d) $N_{\text{bins}} = 10$, $\text{p}^{\text{syst}} = 50\%$

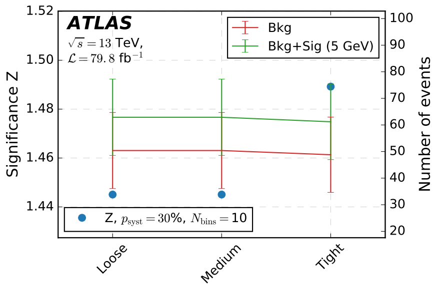
Figure 74: Signal sample with $M_N = 50$ GeV. Optimization study for the variable E_T^{miss}. The plots show the significance and number of background and signal events as a function of the cut on E_T^{miss} for which the x-axis value x_0 represents the requirement $E_T^{\text{miss}} \leq x_0$ GeV.



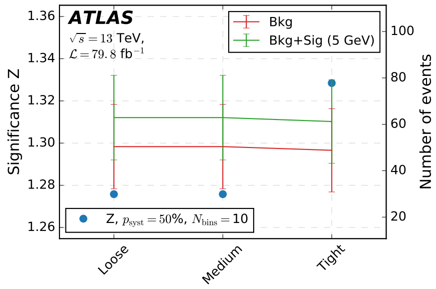
(a) $N_{\text{bins}} = 5$, $\text{p}^{\text{syst}} = 30\%$



(b) $N_{\text{bins}} = 5$, $\text{p}^{\text{syst}} = 50\%$

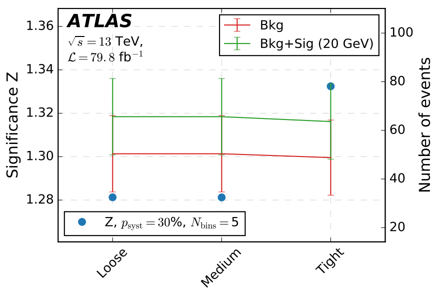


(c) $N_{\text{bins}} = 10$, $\text{p}^{\text{syst}} = 30\%$

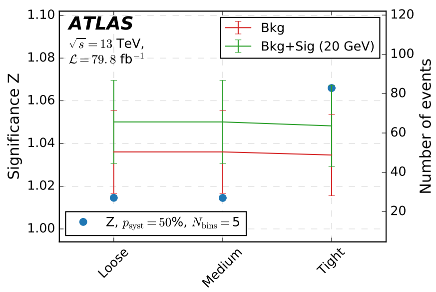


(d) $N_{\text{bins}} = 10$, $\text{p}^{\text{syst}} = 50\%$

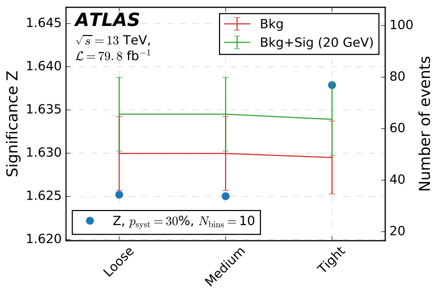
Figure 75: Signal sample with $M_N = 5$ GeV. Optimization study for the identification quality for μ_1. The plots show the significance and number of background and signal events as a function of the cut on the identification quality selection for μ_1.



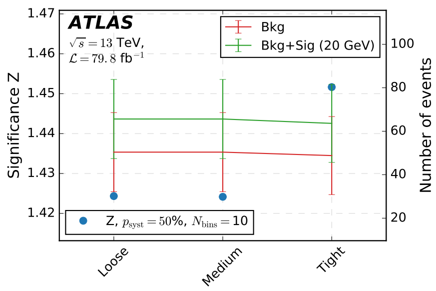
(a) $N_{\text{bins}} = 5$, $\text{p}^{\text{syst}} = 30\%$



(b) $N_{\text{bins}} = 5$, $\text{p}^{\text{syst}} = 50\%$

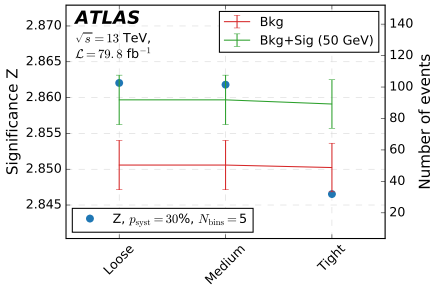


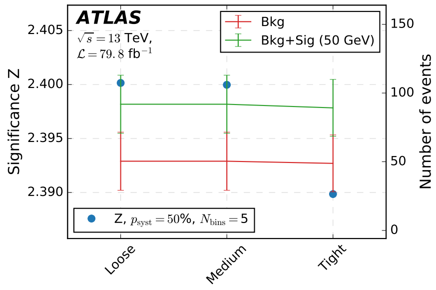
(c) $N_{\text{bins}} = 10$, $\text{p}^{\text{syst}} = 30\%$



(d) $N_{\text{bins}} = 10$, $\text{p}^{\text{syst}} = 50\%$

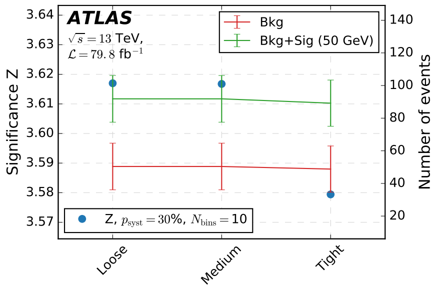
Figure 76: Signal sample with $M_N = 20$ GeV. Optimization study for the identification quality for μ_1. The plots show the significance and number of background and signal events as a function of the cut on the identification quality selection for μ_1.

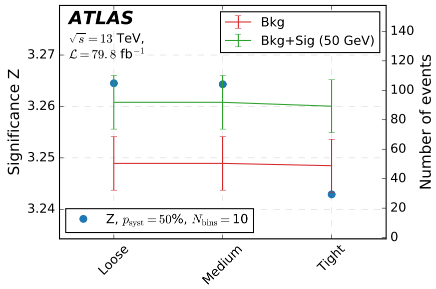




(a) $N_{\mathrm{bins}} = 5$, $\mathrm{p}^{\mathrm{syst}} = 30\%$

(b) $N_{\mathrm{bins}} = 5$, $\mathrm{p}^{\mathrm{syst}} = 50\%$

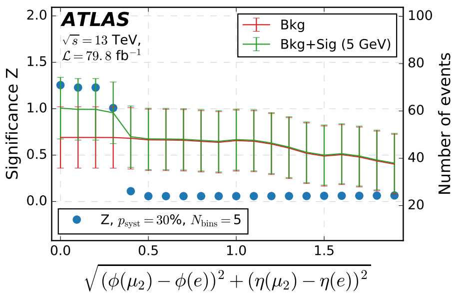


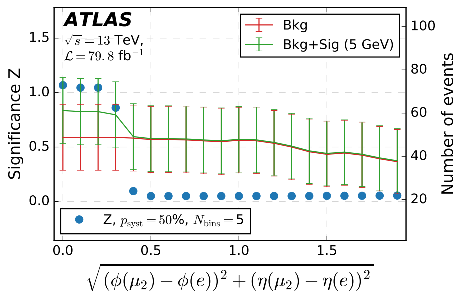


(c) $N_{\mathrm{bins}} = 10$, $\mathrm{p}^{\mathrm{syst}} = 30\%$

(d) $N_{\mathrm{bins}} = 10$, $\mathrm{p}^{\mathrm{syst}} = 50\%$

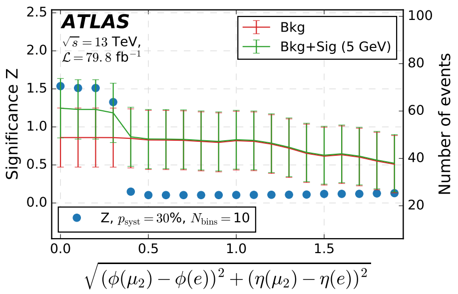
Figure 77: Signal sample with $M_N = 50$ GeV. Optimization study for the identification quality for μ_1. The plots show the significance and number of background and signal events as a function of the cut on the identification quality selection for μ_1.

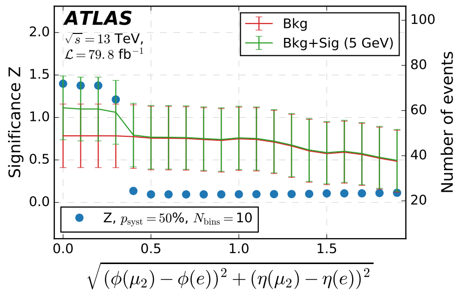




(a) $N_{\mathrm{bins}} = 5$, $\mathrm{p}^{\mathrm{syst}} = 30\%$

(b) $N_{\mathrm{bins}} = 5$, $\mathrm{p}^{\mathrm{syst}} = 50\%$

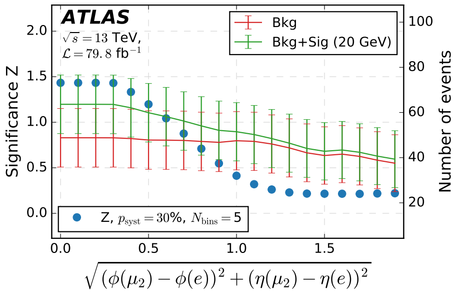




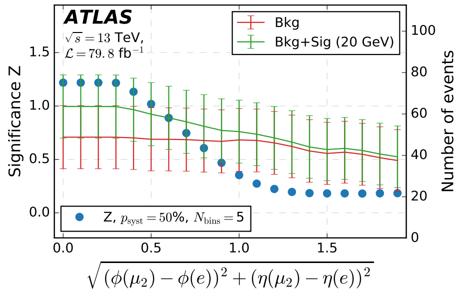
(c) $N_{\mathrm{bins}} = 10$, $\mathrm{p}^{\mathrm{syst}} = 30\%$

(d) $N_{\mathrm{bins}} = 10$, $\mathrm{p}^{\mathrm{syst}} = 50\%$

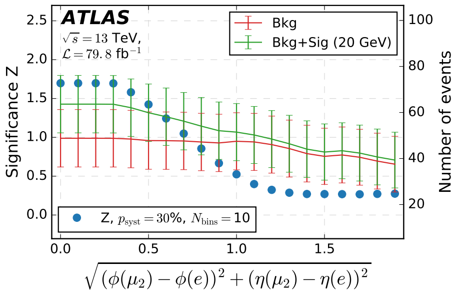
Figure 78: Signal sample with $M_N = 5$ GeV. Optimization study for variable $\Delta R(\mu_2, e)$. The plots show the significance and number of events as a function of the cut on $\Delta R(\mu_2, e)$ for which the x-axis value x_0 represents the requirement $\Delta R(\mu_2, e) \geq x_0$ GeV.



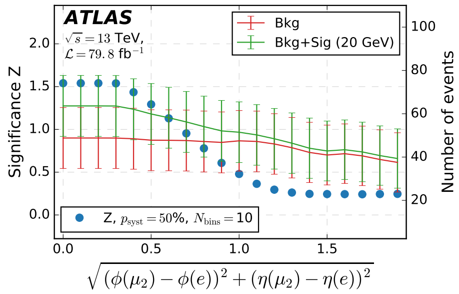
(a) $N_{\rm bins} = 5$, $\mathrm{p}^{\rm syst} = 30\%$



(b) $N_{\rm bins} = 5$, $\mathrm{p}^{\rm syst} = 50\%$

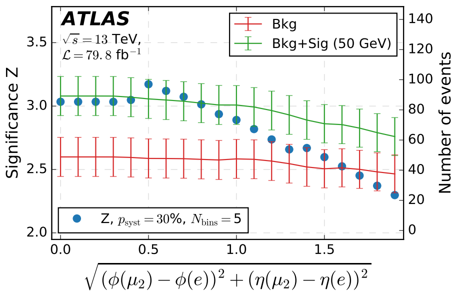


(c) $N_{\rm bins} = 10$, $\mathrm{p}^{\rm syst} = 30\%$

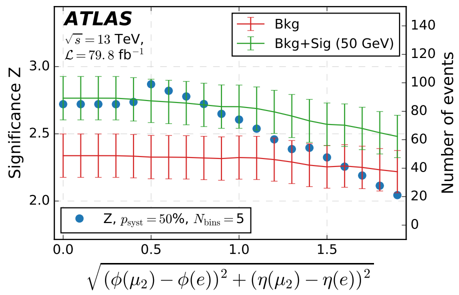


(d) $N_{\rm bins} = 10$, $\mathrm{p}^{\rm syst} = 50\%$

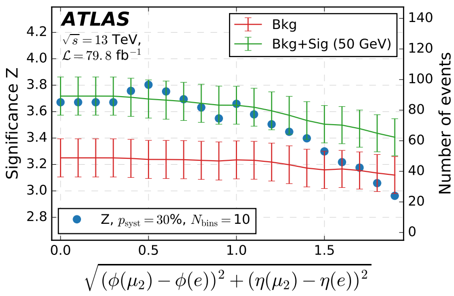
Figure 79: Signal sample with $M_N = 20$ GeV. Optimization study for variable $\Delta R(\mu_2, e)$. The plots show the significance and number of events as a function of the cut on $\Delta R(\mu_2, e)$ for which the x-axis value x_0 represents the requirement $\Delta R(\mu_2, e) \geq x_0$ GeV.



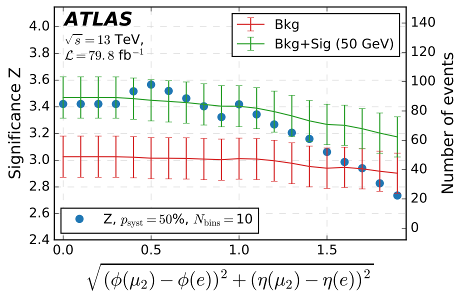
(a) $N_{\rm bins} = 5$, $\mathrm{p}^{\rm syst} = 30\%$



(b) $N_{\rm bins} = 5$, $\mathrm{p}^{\rm syst} = 50\%$

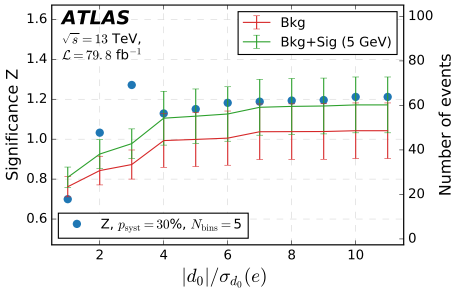


(c) $N_{\rm bins} = 10$, $\mathrm{p}^{\rm syst} = 30\%$

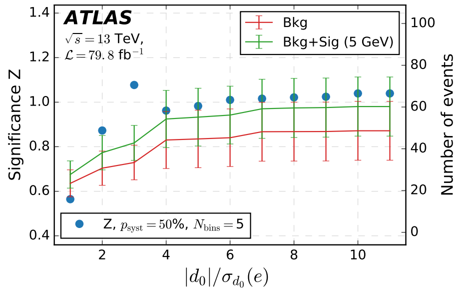


(d) $N_{\rm bins} = 10$, $\mathrm{p}^{\rm syst} = 50\%$

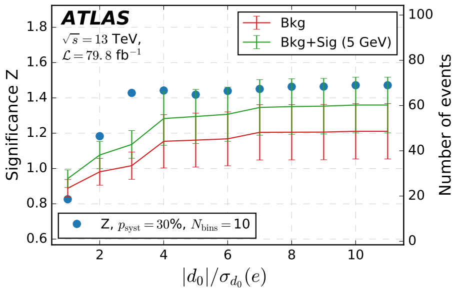
Figure 80: Signal sample with $M_N = 50$ GeV. Optimization study for variable $\Delta R(\mu_2, e)$. The plots show the significance and number of events as a function of the cut on $\Delta R(\mu_2, e)$ for which the x-axis value x_0 represents the requirement $\Delta R(\mu_2, e) \geq x_0$ GeV.



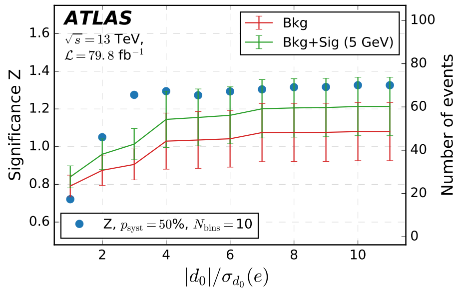
(a) $N_{\text{bins}} = 5$, $\text{p}^{\text{syst}} = 30\%$



(b) $N_{\text{bins}} = 5$, $\text{p}^{\text{syst}} = 50\%$

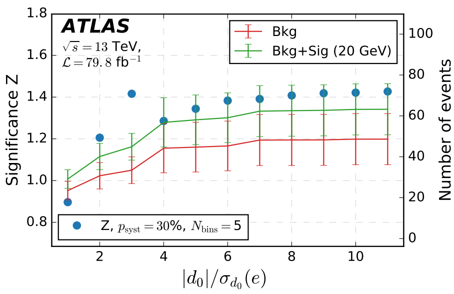


(c) $N_{\text{bins}} = 10$, $\text{p}^{\text{syst}} = 30\%$

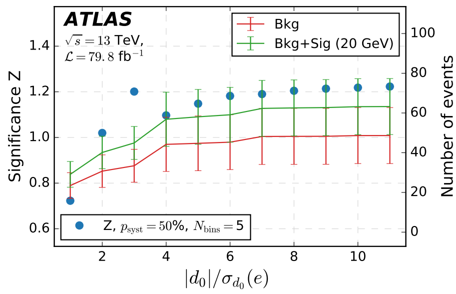


(d) $N_{\text{bins}} = 10$, $\text{p}^{\text{syst}} = 50\%$

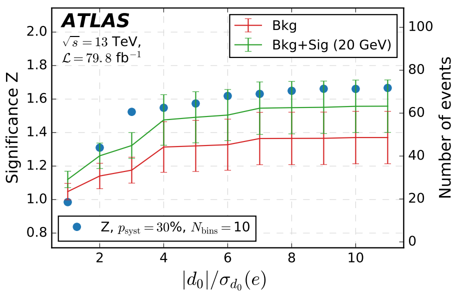
Figure 81: Signal sample with $M_N = 5$ GeV. Optimization study for the variable $|d_0|/\sigma(d_0)(e)$. The plots show the significance and number of events as a function of the cut on the variable $|d_0|/\sigma(d_0)(e)$ for which the x-axis value x_0 represents the requirement $|d_0|/\sigma(d_0)(e) < x_0$.



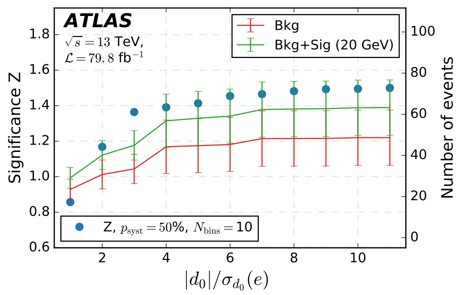
(a) $N_{\text{bins}} = 5$, $\text{p}^{\text{syst}} = 30\%$



(b) $N_{\text{bins}} = 5$, $\text{p}^{\text{syst}} = 50\%$

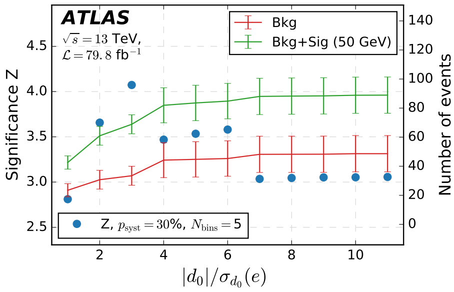


(c) $N_{\text{bins}} = 10$, $\text{p}^{\text{syst}} = 30\%$

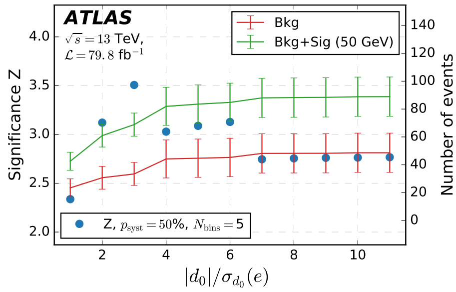


(d) $N_{\text{bins}} = 10$, $\text{p}^{\text{syst}} = 50\%$

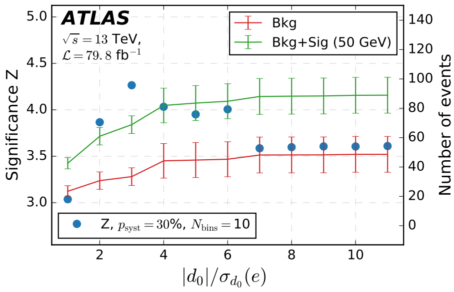
Figure 82: Signal sample with $M_N = 20$ GeV. Optimization study for the variable $|d_0|/\sigma(d_0)(e)$. The plots show the significance and number of events as a function of the cut on the variable $|d_0|/\sigma(d_0)(e)$ for which the x-axis value x_0 represents the requirement $|d_0|/\sigma(d_0)(e) < x_0$.



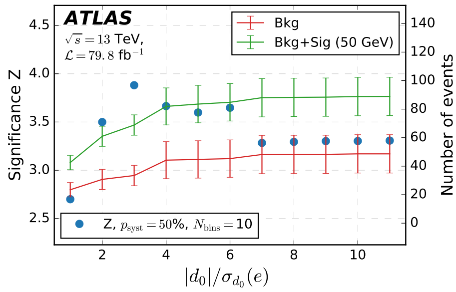
(a) $N_{\text{bins}} = 5$, $\text{p}^{\text{syst}} = 30\%$



(b) $N_{\text{bins}} = 5$, $\text{p}^{\text{syst}} = 50\%$

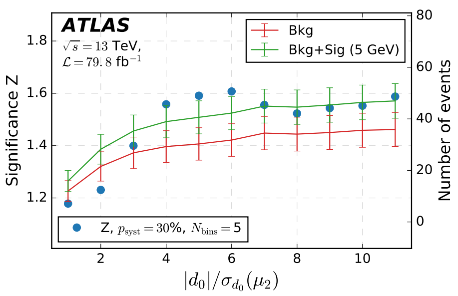


(c) $N_{\text{bins}} = 10$, $\text{p}^{\text{syst}} = 30\%$

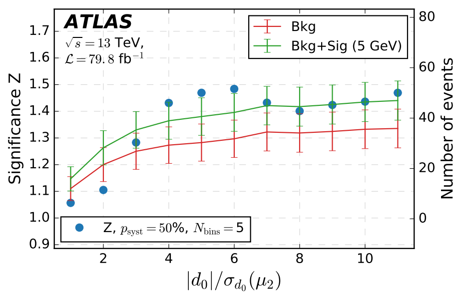


(d) $N_{\text{bins}} = 10$, $\text{p}^{\text{syst}} = 50\%$

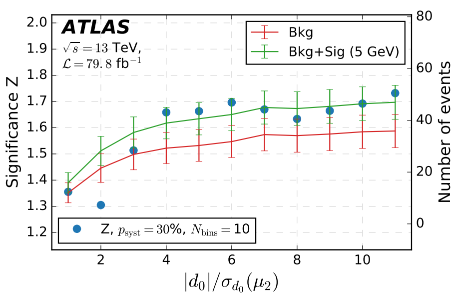
Figure 83: Signal sample with $M_N = 50$ GeV. Optimization study for the variable $|d_0|/\sigma(d_0)(e)$. The plots show the significance and number of events as a function of the cut on the variable $|d_0|/\sigma(d_0)(e)$ for which the x-axis value x_0 represents the requirement $|d_0|/\sigma(d_0)(e) < x_0$.



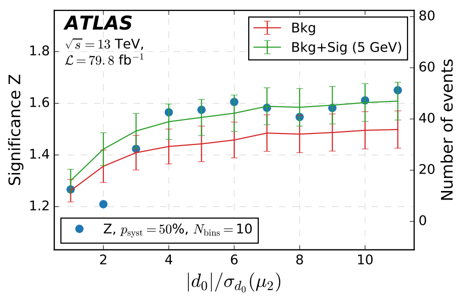
(a) $N_{\text{bins}} = 5$, $\text{p}^{\text{syst}} = 30\%$



(b) $N_{\text{bins}} = 5$, $\text{p}^{\text{syst}} = 50\%$

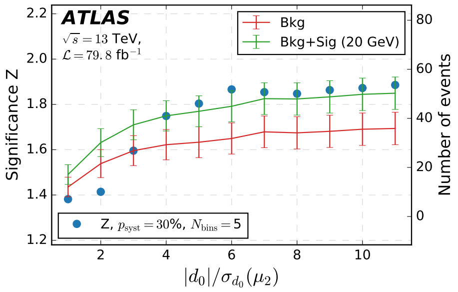


(c) $N_{\text{bins}} = 10$, $\text{p}^{\text{syst}} = 30\%$

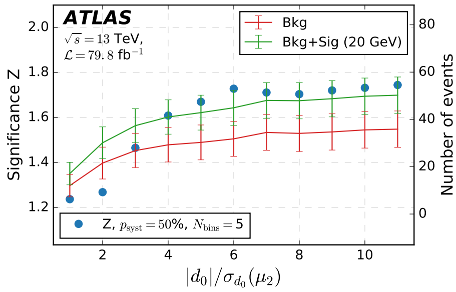


(d) $N_{\text{bins}} = 10$, $\text{p}^{\text{syst}} = 50\%$

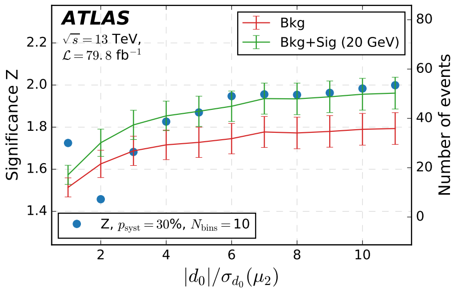
Figure 84: Signal sample with $M_N = 5$ GeV. Optimization study for the variable $|d_0|/\sigma(d_0)(\mu_2)$. The plots show the significance and number of events as a function of the cut on the variable $|d_0|/\sigma(d_0)(\mu_2)$ for which the x-axis value x_0 represents the requirement $|d_0|/\sigma(d_0)(\mu_2) < x_0$.



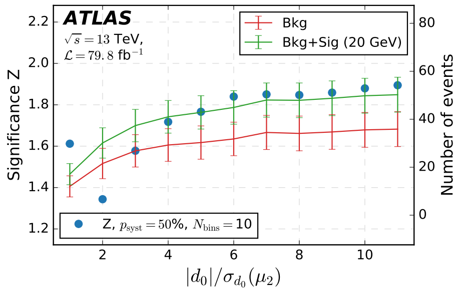
(a) $N_{\text{bins}} = 5$, $\text{p}^{\text{syst}} = 30\%$



(b) $N_{\text{bins}} = 5$, $\text{p}^{\text{syst}} = 50\%$

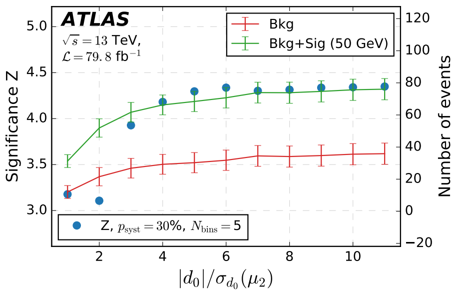


(c) $N_{\text{bins}} = 10$, $\text{p}^{\text{syst}} = 30\%$

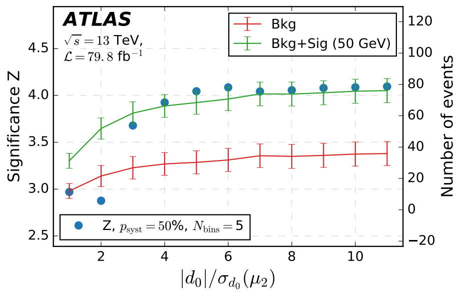


(d) $N_{\text{bins}} = 10$, $\text{p}^{\text{syst}} = 50\%$

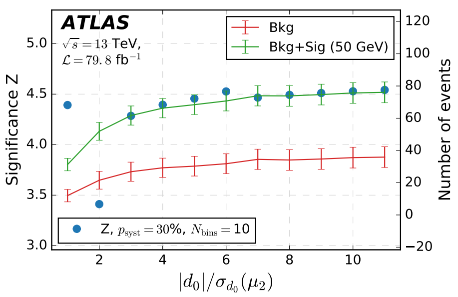
Figure 85: Signal sample with $M_N = 20$ GeV. Optimization study for the variable $|d_0|/\sigma(d_0)(\mu_2)$. The plots show the significance and number of events as a function of the cut on the variable $|d_0|/\sigma(d_0)(\mu_2)$ for which the x-axis value x_0 represents the requirement $|d_0|/\sigma(d_0)(\mu_2) < x_0$.



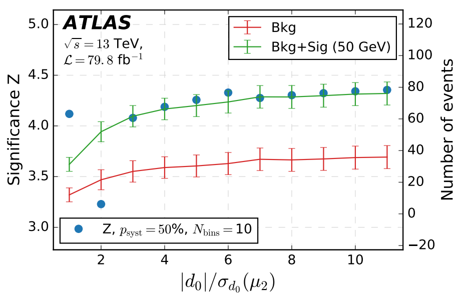
(a) $N_{\text{bins}} = 5$, $\text{p}^{\text{syst}} = 30\%$



(b) $N_{\text{bins}} = 5$, $\text{p}^{\text{syst}} = 50\%$

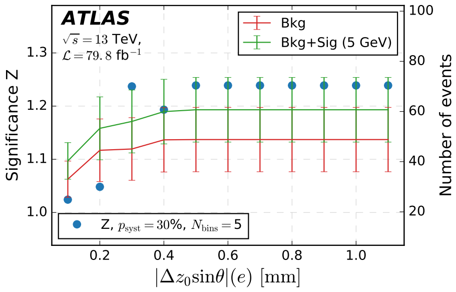


(c) $N_{\text{bins}} = 10$, $\text{p}^{\text{syst}} = 30\%$

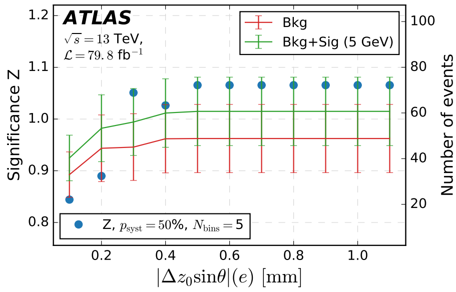


(d) $N_{\text{bins}} = 10$, $\text{p}^{\text{syst}} = 50\%$

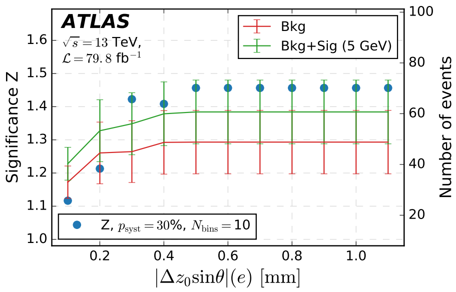
Figure 86: Signal sample with $M_N = 50$ GeV. Optimization study for the variable $|d_0|/\sigma(d_0)(\mu_2)$. The plots show the significance and number of events as a function of the cut on the variable $|d_0|/\sigma(d_0)(\mu_2)$ for which the x-axis value x_0 represents the requirement $|d_0|/\sigma(d_0)(\mu_2) < x_0$.



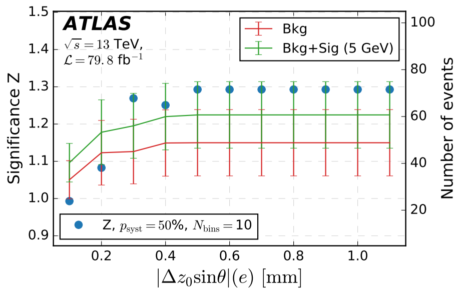
(a) $N_{\rm bins} = 5$, $\rm p^{syst} = 30\%$



(b) $N_{\rm bins} = 5$, $\rm p^{syst} = 50\%$

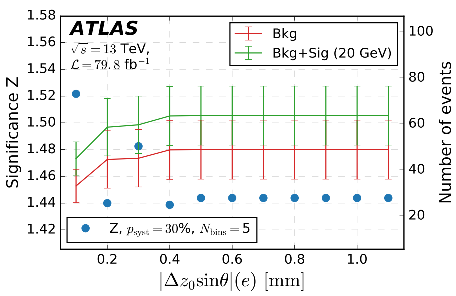


(c) $N_{\rm bins} = 10$, $\rm p^{syst} = 30\%$

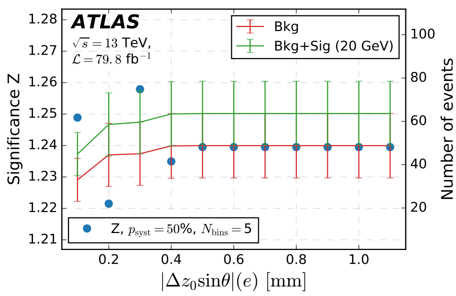


(d) $N_{\rm bins} = 10$, $\rm p^{syst} = 50\%$

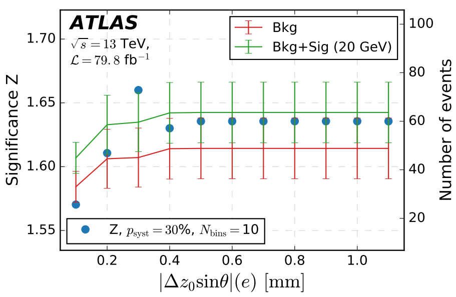
Figure 87: Signal sample with $M_N = 5$ GeV. Optimization study for the variable $|\Delta z_0 \sin(\theta)|(e)$. The plots show the significance and number of events as a function of the cut on the variable $|\Delta z_0 \sin(\theta)|(e)$ for which the x-axis value x_0 represents the requirement $|\Delta z_0 \sin(\theta)|(e) < x_0$.



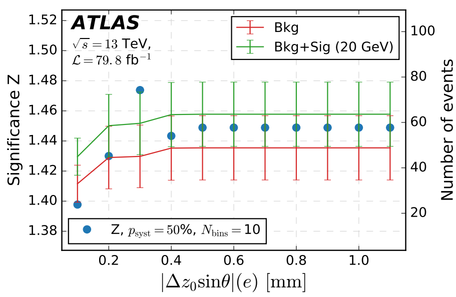
(a) $N_{\rm bins} = 5$, $\rm p^{syst} = 30\%$



(b) $N_{\rm bins} = 5$, $\rm p^{syst} = 50\%$

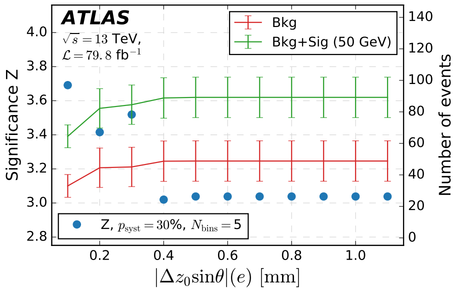


(c) $N_{\rm bins} = 10$, $\rm p^{syst} = 30\%$

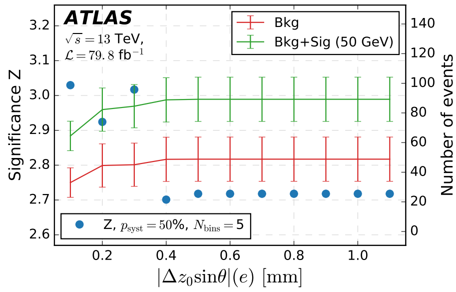


(d) $N_{\rm bins} = 10$, $\rm p^{syst} = 50\%$

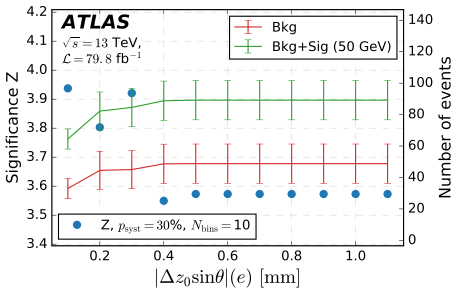
Figure 88: Signal sample with $M_N = 20$ GeV. Optimization study for the variable $|\Delta z_0 \sin(\theta)|(e)$. The plots show the significance and number of events as a function of the cut on the variable $|\Delta z_0 \sin(\theta)|(e)$ for which the x-axis value x_0 represents the requirement $|\Delta z_0 \sin(\theta)|(e) < x_0$.



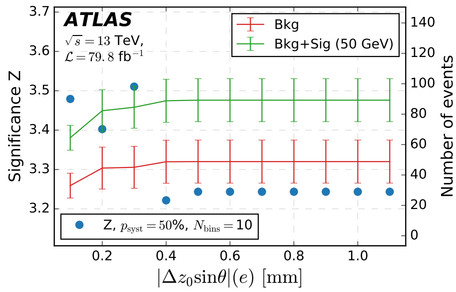
(a) $N_{\text{bins}} = 5$, $\text{p}^{\text{syst}} = 30\%$



(b) $N_{\text{bins}} = 5$, $\text{p}^{\text{syst}} = 50\%$

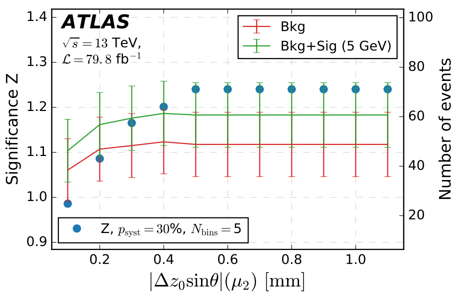


(c) $N_{\text{bins}} = 10$, $\text{p}^{\text{syst}} = 30\%$

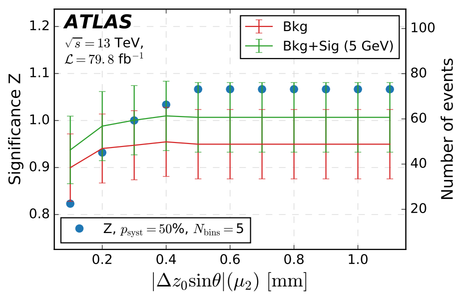


(d) $N_{\text{bins}} = 10$, $\text{p}^{\text{syst}} = 50\%$

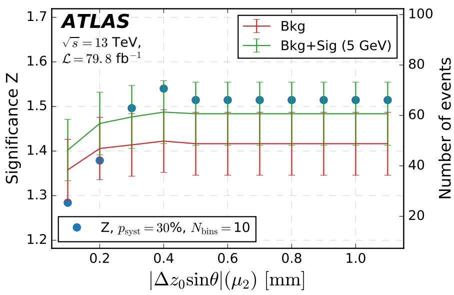
Figure 89: Signal sample with $M_N = 50$ GeV. Optimization study for the variable $|\Delta z_0 \sin(\theta)|(e)$. The plots show the significance and number of events as a function of the cut on the variable $|\Delta z_0 \sin(\theta)|(e)$ for which the x-axis value x_0 represents the requirement $|\Delta z_0 \sin(\theta)|(e) < x_0$.



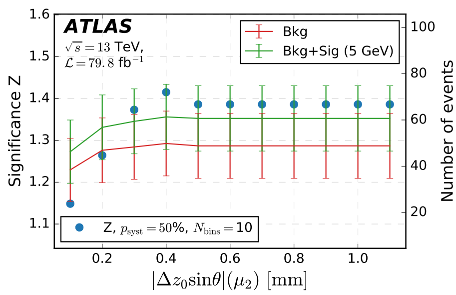
(a) $N_{\text{bins}} = 5$, $\text{p}^{\text{syst}} = 30\%$



(b) $N_{\text{bins}} = 5$, $\text{p}^{\text{syst}} = 50\%$

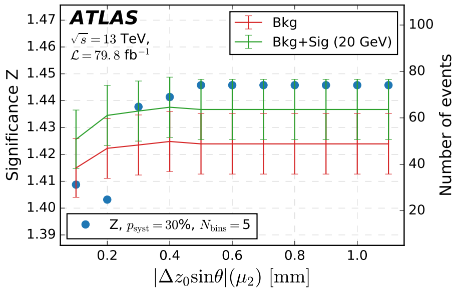


(c) $N_{\text{bins}} = 10$, $\text{p}^{\text{syst}} = 30\%$

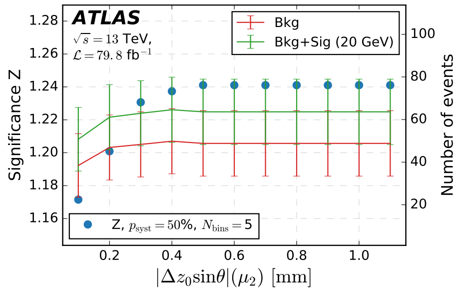


(d) $N_{\text{bins}} = 10$, $\text{p}^{\text{syst}} = 50\%$

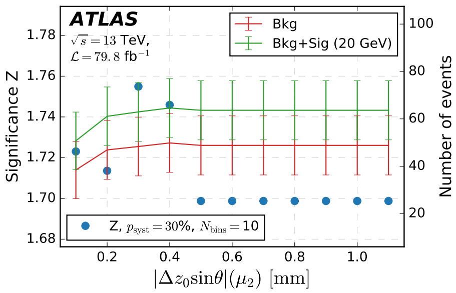
Figure 90: Signal sample with $M_N = 5$ GeV. Optimization study for the variable $|\Delta z_0 \sin(\theta)|(\mu_2)$. The plots show the significance and number of events as a function of the cut on the variable $|\Delta z_0 \sin(\theta)|(\mu_2)$ for which the x-axis value x_0 represents the requirement $|\Delta z_0 \sin(\theta)|(\mu_2) < x_0$.



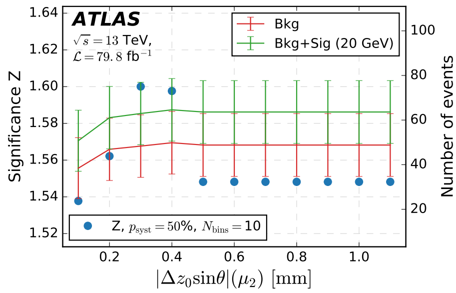
(a) $N_{\mathrm{bins}} = 5$, $\mathrm{p}^{\mathrm{syst}} = 30\%$



(b) $N_{\mathrm{bins}} = 5$, $\mathrm{p}^{\mathrm{syst}} = 50\%$

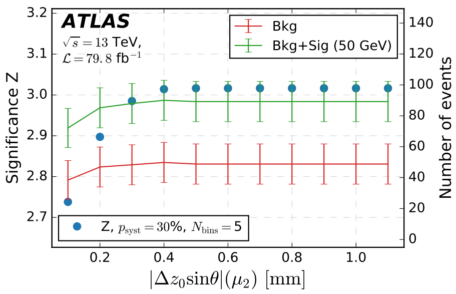


(c) $N_{\mathrm{bins}} = 10$, $\mathrm{p}^{\mathrm{syst}} = 30\%$

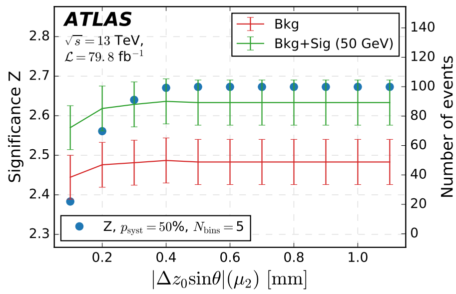


(d) $N_{\mathrm{bins}} = 10$, $\mathrm{p}^{\mathrm{syst}} = 50\%$

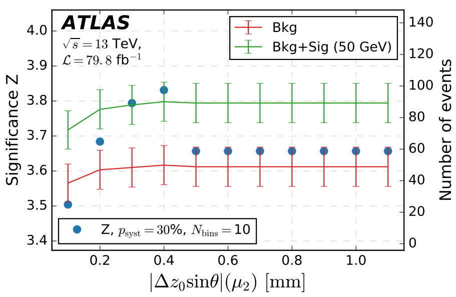
Figure 91: Signal sample with $M_N = 20$ GeV. Optimization study for the variable $|\Delta z_0 \sin(\theta)|(\mu_2)$. The plots show the significance and number of events as a function of the cut on the variable $|\Delta z_0 \sin(\theta)|(\mu_2)$ for which the x-axis value x_0 represents the requirement $|\Delta z_0 \sin(\theta)|(\mu_2) < x_0$.



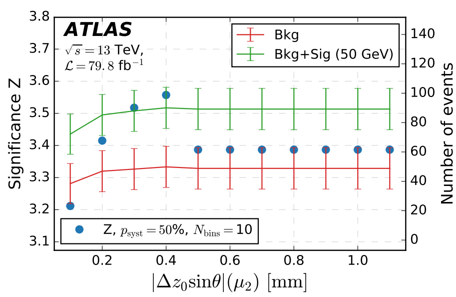
(a) $N_{\mathrm{bins}} = 5$, $\mathrm{p}^{\mathrm{syst}} = 30\%$



(b) $N_{\mathrm{bins}} = 5$, $\mathrm{p}^{\mathrm{syst}} = 50\%$



(c) $N_{\mathrm{bins}} = 10$, $\mathrm{p}^{\mathrm{syst}} = 30\%$



(d) $N_{\mathrm{bins}} = 10$, $\mathrm{p}^{\mathrm{syst}} = 50\%$

Figure 92: Signal sample with $M_N = 50$ GeV. Optimization study for the variable $|\Delta z_0 \sin(\theta)|(\mu_2)$. The plots show the significance and number of events as a function of the cut on the variable $|\Delta z_0 \sin(\theta)|(\mu_2)$ for which the x-axis value x_0 represents the requirement $|\Delta z_0 \sin(\theta)|(\mu_2) < x_0$.

D Signal region cutflows

The cutflows for the signal samples and background processes are presented in this section. A "cut" refers to a signal region requirements. The cutflows show the number of events for a given process at each selection stage. The signal region requirements are listed in Tab. 65. The cutflows start at cut number 10, since the previous requirements are applied doing the object definition or are used as requirements in the MC derivation [65] applied.

The cutflow for the signal samples is shown in Tab. 66. Table 67 and 68 show cutflows for the background processes. The sum of all the background processes applied is listed in Tab. 68.

Preselection requirements	
1.	Two `loose` and `combined` muons
2.	One `LooseAndBLayerLLH` electron
3.	$\mu^{\pm}\mu^{\pm}e^{\mp}$
4.	$\lvert\eta(\mu)\rvert < 2.5$ and $\lvert\eta(e)\rvert < 2.47$ except $1.37 < \lvert\eta(e)\rvert < 1.52$
5.	$\mu : p_T \geq 4$ GeV
6.	$e : p_T \geq 7$ GeV (2015), $p_T \geq 4.5$ GeV (2016+2017)
7.	$\mu_1 : \lvert d_0\rvert/\sigma(d_0) < 3$ and $\lvert\Delta z_0 \sin\theta\rvert < 0.5$ mm
8.	$\mu_2 + e : \lvert\Delta z_0 \sin\theta\rvert < 1.0$ mm
9.	Trigger: `HLT_mu18_mu8noL1` (2015), `HLT_mu22_mu8noL1` (2016+2017)
10.	Trigger matching and offline p_T requirements
11.	$\mu_2 : p_T < 50$ GeV
12.	$e : p_T < 50$ GeV
13.	$40\,\text{GeV} < M(\mu_1, \mu_2, e) < 90\,\text{GeV}$
14.	$M(\mu_2, e) < 50$ GeV
Optimization study requirements	
15.	$p_T(\mu_1) > 23$ GeV
16.	$p_T(\mu_2) > 10$ GeV
17.	$p_T(e) > 10$ GeV
18.	$N_{\text{b-jets}} = 0$
19.	Isolation for μ_1: `FixedCutTightTrackOnly`
20.	Isolation for μ_2: `FixedCutLoose`
21.	Isolation for e: `Gradient`
22.	$E_T^{\text{miss}} < 60$ GeV
23.	Identification selection for μ_1: `Tight`
24.	$\Delta R(\mu_2, e) > 0.2$
25.	$\lvert d_0\rvert/\sigma(d_0)(\mu_2) < 10$ and $\lvert\Delta z_0 \sin(\theta)\rvert(\mu_2) < 0.5$ mm
26.	$\lvert d_0\rvert/\sigma(d_0)(e) < 10$ and $\lvert\Delta z_0 \sin(\theta)\rvert(e) < 0.5$ mm

Table 65: The final list of signal region requirements.

Cut	5 GeV	10 GeV	20 GeV	30 GeV	50 GeV
10	721±29	31±6	37±6	117±11	80±9
11	719±29	31±6	37±6	117±11	80±9
12	714±28	31±6	36±6	116±11	79±9
13	699±28	30±6	35±6	113±11	74±9
14	695±28	30±6	35±6	112±11	74±9
15	694±28	30±6	35±6	112±11	72±9
16	381±21	19±4	23±5	74±9	56±8
17	340±20	17±4	20±5	66±8	52±7
18	335±20	17±4	20±5	64±8	51±7
19	305±19	15±4	19±4	60±8	47±7
20	13±4	9±3	16±4	55±7	44±7
21	12±4	9±3	15±4	51±7	40±6
22	12±4	9±3	15±4	50±7	40±6
23	12±4	8±3	14±4	49±7	39±6
24	12±4	8±3	14±4	49±7	39±6
25	11±3	6±3	14±4	47±7	38±6
26	11±3	6±3	14±4	47±7	38±6

Table 66: The number of events in each signal sample with M_N = 5, 10, 20, 30 and 50 GeV at each step of the SR requirement list, see Tab. 65. Scale factors are applied for jets, the leading muon and the electron.

Cut	$t\bar{t}$	W+jets	Z+jets	t and Wt
10	7115±95	1382±250	180±39	559±27
11	7010±95	1380±250	179±39	554±27
12	4595±77	1358±250	169±39	393±22
13	1914±50	737±221	136±27	175±15
14	1824±48	725±221	137±27	171±15
15	1817±48	722±221	136±27	171±15
16	1089±37	78±43	60±18	89±11
17	911±34	85±38	51±16	76±10
18	367±22	78±38	46±16	54±8
19	320±20	73±38	42±15	48±8
20	54±8	10±6	25±12	8±3
21	40±7	0.5±0.9	20±12	4±2
22	18±5	0.7±0.9	20±12	2±2
23	18±5	0.7±0.9	19±12	2±2
24	18±5	0.7±0.9	19±12	2±2
25	16±5	0.7±0.9	10±6	2±2
26	16±5	0.7±0.9	10±6	2±2

Table 67: The number of events for background processes at each step of the SR requirement list, see Tab. 65. Scale factors are applied for jets, the leading muon and the electron.

Cut	Triboson	$t\bar{t}V$	$V\gamma^*$	Diboson	$\sum$ Bkg
10	21±5	58±8	30±9	110±12	9454±272
11	17±4	46±7	30±9	104±10	9321±272
12	11±3	28±5	28±9	92±10	6675±266
13	4±2	7±3	11±5	40±6	3023±229
14	3±2	7±3	11±5	38±6	2915±228
15	3±2	7±3	11±5	38±6	2905±228
16	2±2	4±2	5±4	20±5	1347±61
17	2±2	4±2	5±3	17±4	1151±55
18	2±2	0.8±0.9	5±3	16±4	568±48
19	2±1	0.7±0.8	3±3	15±4	504±46
20	2±1	0.4±0.6	0.7±1.1	12±3	111±17
21	2±1	0.3±0.5	0.7±1.1	10±3	78±14
22	0.7±0.9	0.1±0.3	0.7±1.1	7±3	49±13
23	0.7±0.8	0.1±0.3	0.7±1.1	7±3	48±13
24	0.7±0.8	0.1±0.3	0.7±1.1	7±3	48±13
25	0.7±0.8	0.1±0.3	0.0±0.0	6±3	35±8
26	0.7±0.8	0.1±0.3	0.0±0.0	6±3	35±8

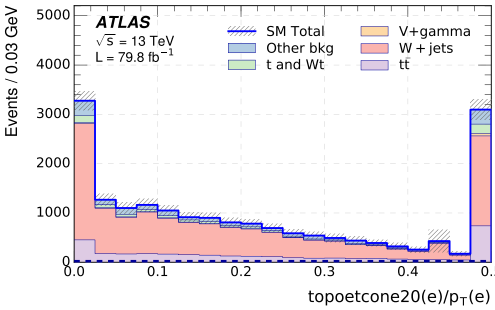
Table 68: The number of events for background processes at each step of the SR requirement list, see Tab. 65. The total number of background events is also listed which includes both the background processes shown in this table and in table 67. Scale factors are applied for jets, the leading muon and the electron.

E Fake electron comparison

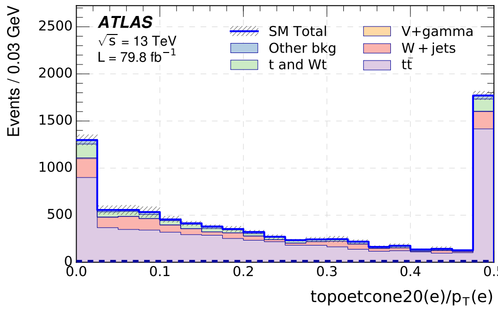
The fake electrons in the SR and electron measurement region are mainly of the three following types: heavy-flavour decays labeled “non-isolated”; electron fakes which originate from light flavour jets labeled “hadrons”; and electron fakes from photon conversion and decays of light mesons labeled “bkg lep”. In this section, the properties of the three fake electron types are compared. The considered variables are: topoetcone20/$p_T(e_{\mathrm{FF}})$, ptvarcone20/$p_T(e_{\mathrm{FF}})$, $p_T(e_{\mathrm{FF}})$, $|\eta|(e_{\mathrm{FF}})$, $|d_0|/\sigma(d_0)(e_{\mathrm{FF}})$; $|\Delta z_0 \sin(\theta)|(e_{\mathrm{FF}})$ and E_T^{miss}. The distributions are showing events in the electron measurement region. Table 69 provides an overview of the figures used in the study of the three fake electron compositions and a summary of how the kinematics of the three fake lepton types differ.

Figure	Distribution	Comparison of distributions
Fig. 93	topoetcone20/$p_T(e_{\mathrm{FF}})$	“Non-isolated” fake electrons tend to be most isolated in the calorimeter and “hadrons” the least isolated.
Fig. 94	ptvarcone20/$p_T(e_{\mathrm{FF}})$	Same tendency in the inner detector as for the calorimeter isolation above.
Fig. 95	$p_T(e_{\mathrm{FF}})$	In all cases, the majority of fake electrons have small p_T. “Non-isolated” fake electrons are most likely to have a higher p_T, especially in the region with $N_{\text{b-jets}} \geq 1$.
Fig. 96	$\|\eta\|(e_{\mathrm{FF}})$	The number of “non-isolated” fake electrons is slightly decreasing as a function of $\|\eta\|$ whereas the number of “hadrons” fake electrons is slightly increasing.
Fig. 97	$\|d_0\|/\sigma(d_0)(e_{\mathrm{FF}})$	“Non-isolated” fake electrons tend to have a higher $\|d_0\|/\sigma(d_0)$. Almost all fake electrons of type “hadrons” have $\|d_0\|/\sigma(d_0) < 2$.
Fig. 98	$\|\Delta z_0 \sin(\theta)\|(e_{\mathrm{FF}})$	“Non-isolated” fake electrons have the highest $\|\Delta z_0 \sin(\theta)\|$ in average. “Hadron” fake leptons tend to have the lowest.
Fig. 99	E_T^{miss}	The shape of the distributions are very similar for the three types of fake leptons.
Fig. 100	topoetcone20/$p_T(e_{\mathrm{FF}})$	Requirements for e_{FF}: 10 GeV $< p_T <$ 15 GeV, $\|d_0\|/\sigma(d_0) < 1$, $\|\Delta z_0 \sin(\theta)\| < 0.1$, and $\|\eta\| < 1.0$. Still see the same tendencies as in Fig. 93.
Fig. 101	ptvarcone20/$p_T(e_{\mathrm{FF}})$	Requirements for e_{FF}: 10 GeV $< p_T <$ 15 GeV, $\|d_0\|/\sigma(d_0) < 1$, $\|\Delta z_0 \sin(\theta)\| < 0.1$, and $\|\eta\| < 1.0$. Still see the same tendencies as in Fig. 94.

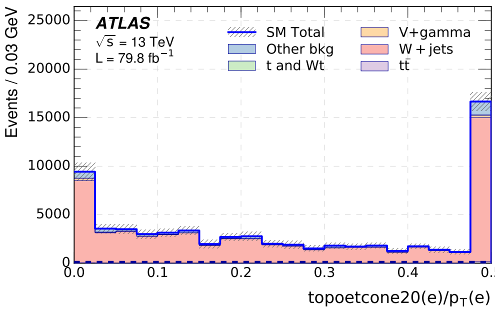
Table 69: Study of the isolation and kinematics of the three types of fake electrons: “non-isolated”, “hadrons” and “bkg lep”, which are defined using truth matching, see Sec. 6.3.



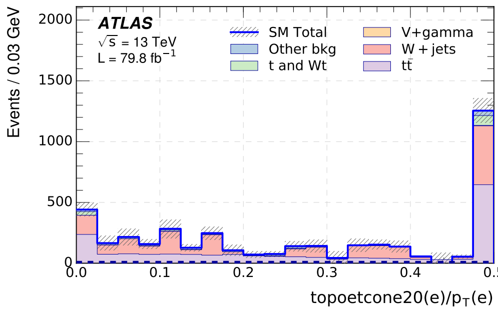
(a) Electron MR with $N_{\text{b-jets}} = 0$.
Truth matching requirement: "Non-isolated" e_{FF}.



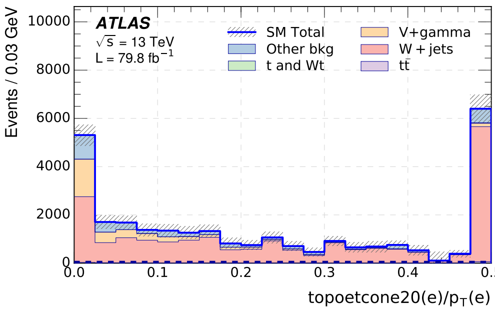
(b) Electron MR with $N_{\text{b-jets}} \geq 1$.
Truth matching requirement: "Non-isolated" e_{FF}.



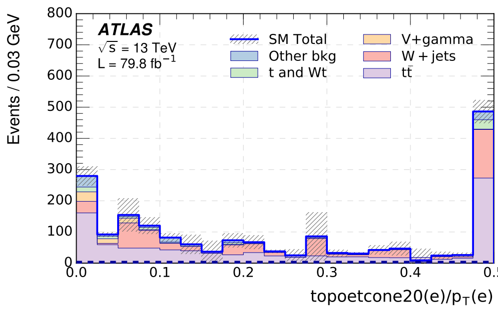
(c) Electron MR with $N_{\text{b-jets}} = 0$.
Truth matching requirement: "Hadron" e_{FF}.



(d) Electron MR with $N_{\text{b-jets}} \geq 1$.
Truth matching requirement: "Hadron" e_{FF}.

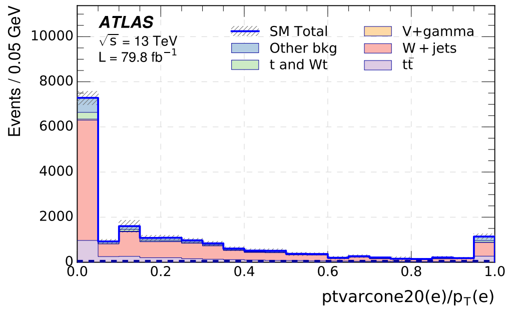


(e) Electron MR with $N_{\text{b-jets}} = 0$.
Truth matching requirement: "Bkg lep" e_{FF}.

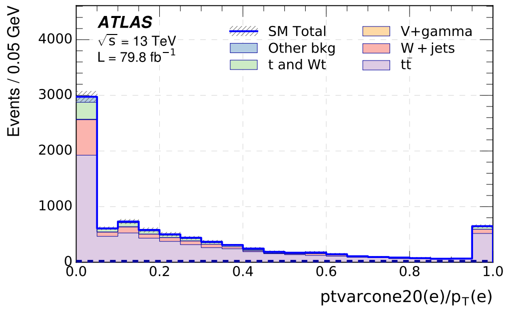


(f) Electron MR with $N_{\text{b-jets}} \geq 1$.
Truth matching requirement: "Bkg lep" e_{FF}.

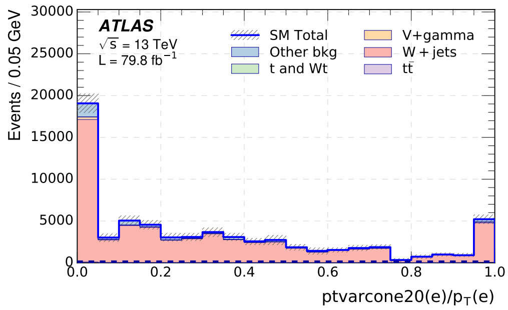
Figure 93: Distribution of the isolation variable: topoetcone20/$p_T(e_{\text{FF}})$. All electron measurement region (MR) requirements are applied except for the requirement that e_{FF} satisfies the `Gradient` isolation selection, see Tab. 15. Truth matching is applied for the fake factor electron in each event and required to be: "non-isolated", "hadron", or "bkg lep". The blue line indicates the sum of the SM background processes and the dashed area the corresponding statistical uncertainty. The last bin in the histogram includes all events with values larger than the range of the histogram.



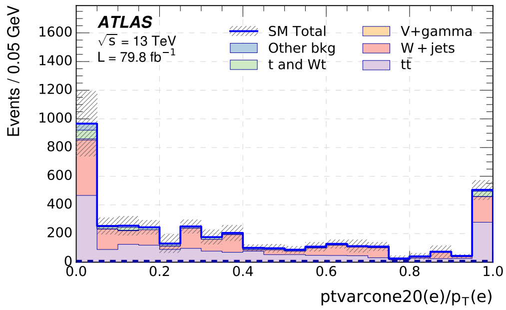
(a) Electron MR with $N_{\text{b-jets}} = 0$.
Truth matching requirement: "Non-isolated" e_{FF}.



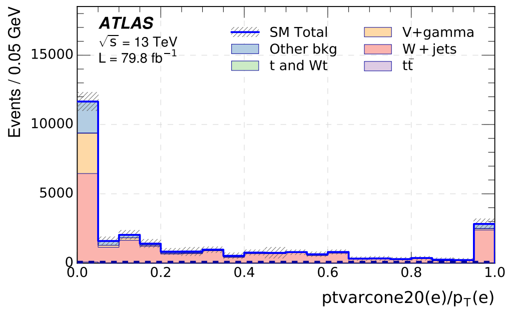
(b) Electron MR with $N_{\text{b-jets}} \geq 1$.
Truth matching requirement: "Non-isolated" e_{FF}.



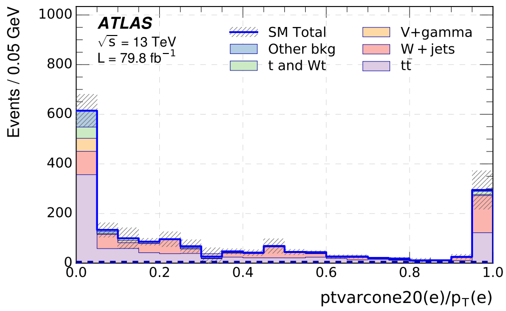
(c) Electron MR with $N_{\text{b-jets}} = 0$.
Truth matching requirement: "Hadron" e_{FF}.



(d) Electron MR with $N_{\text{b-jets}} \geq 1$.
Truth matching requirement: "Hadron" e_{FF}.

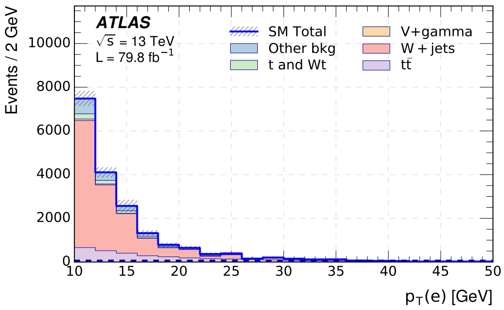


(e) Electron MR with $N_{\text{b-jets}} = 0$.
Truth matching requirement: "Bkg lep" e_{FF}.

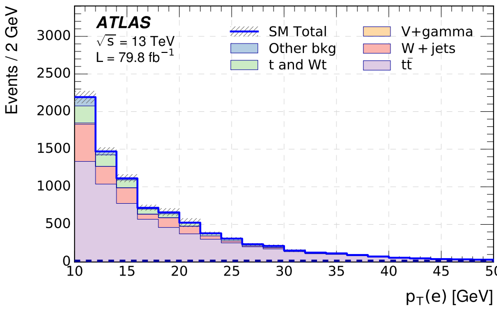


(f) Electron MR with $N_{\text{b-jets}} \geq 1$.
Truth matching requirement: "Bkg lep" e_{FF}.

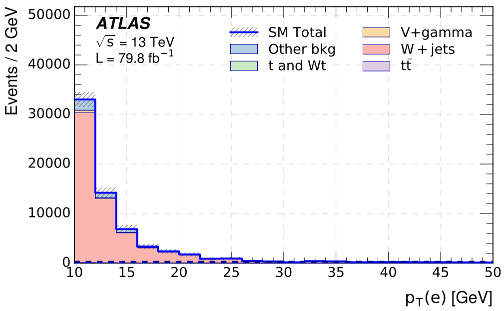
Figure 94: Distribution of the isolation variable: ptvarcone20/$p_T(e_{\text{FF}})$. All electron MR requirements are applied except for the requirement that e_{FF} satisfies the `Gradient` isolation selection, see Tab. 15. Truth matching is applied for the fake factor electron in each event and required to be: "non-isolated", "hadron", or "bkg lep". The blue line indicates the sum of the SM background processes and the dashed area the corresponding statistical uncertainty. The last bin in the histogram includes all events with values larger than the range of the histogram.



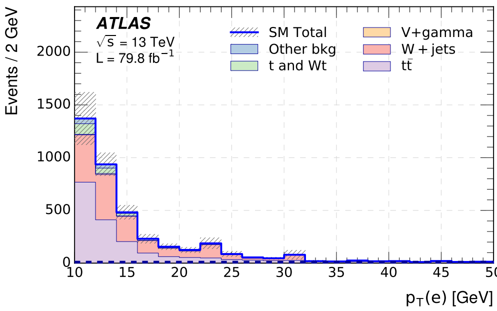
(a) Electron MR with $N_{\text{b-jets}} = 0$.
Truth matching requirement: "Non-isolated" e_{FF}.



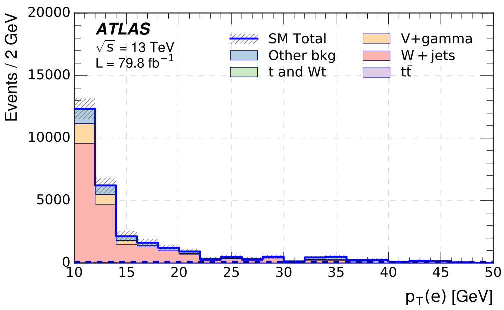
(b) Electron MR with $N_{\text{b-jets}} \geq 1$.
Truth matching requirement: "Non-isolated" e_{FF}.



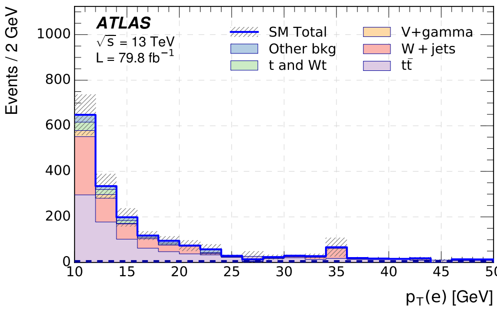
(c) Electron MR with $N_{\text{b-jets}} = 0$.
Truth matching requirement: "Hadron" e_{FF}.



(d) Electron MR with $N_{\text{b-jets}} \geq 1$.
Truth matching requirement: "Hadron" e_{FF}.

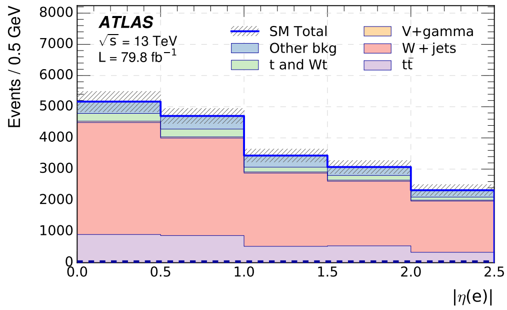


(e) Electron MR with $N_{\text{b-jets}} = 0$.
Truth matching requirement: "Bkg lep" e_{FF}.

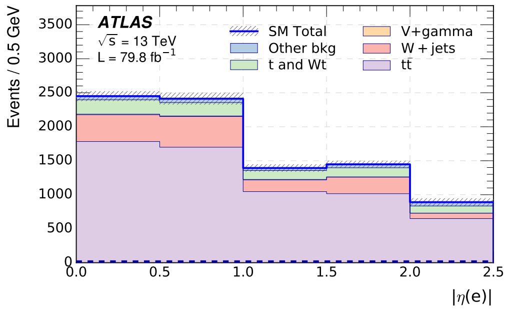


(f) Electron MR with $N_{\text{b-jets}} \geq 1$.
Truth matching requirement: "Bkg lep" e_{FF}.

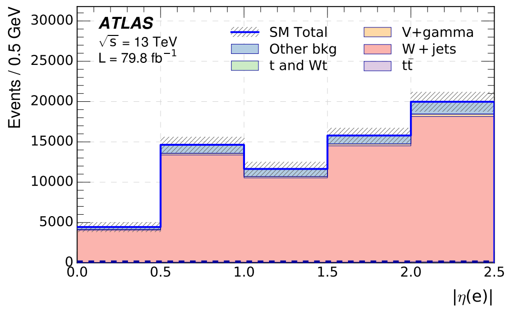
Figure 95: Distribution of the transverse momentum: $p_T(e_{\text{FF}})$. All electron measurement region (MR) requirements are applied except for the requirement that e_{FF} satisfies the `Gradient` isolation selection, see Tab. 15. Truth matching is applied for the fake factor electron in each event and required to be: "non-isolated", "hadron", or "bkg lep". The blue line indicates the sum of the SM background processes and the dashed area the corresponding statistical uncertainty. The last bin in the histogram includes all events with values larger than the range of the histogram.



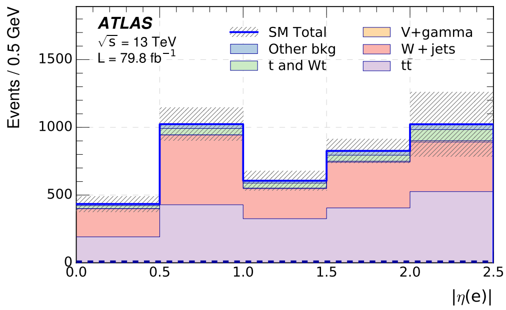
(a) Electron MR with $N_{\text{b-jets}} = 0$.
Truth matching requirement: "Non-isolated" e_{FF}.



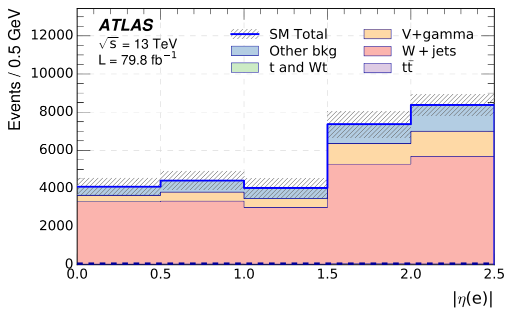
(b) Electron MR with $N_{\text{b-jets}} \geq 1$.
Truth matching requirement: "Non-isolated" e_{FF}.



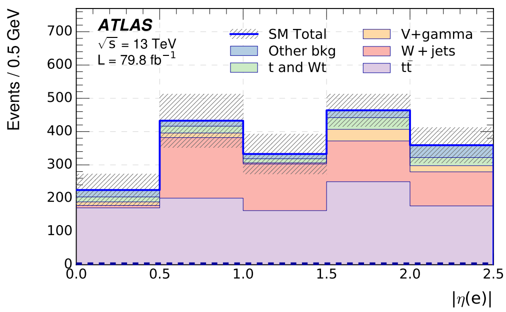
(c) Electron MR with $N_{\text{b-jets}} = 0$.
Truth matching requirement: "Hadron" e_{FF}.



(d) Electron MR with $N_{\text{b-jets}} \geq 1$.
Truth matching requirement: "Hadron" e_{FF}.

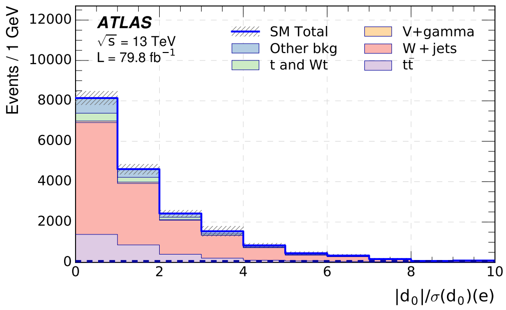


(e) Electron MR with $N_{\text{b-jets}} = 0$.
Truth matching requirement: "Bkg lep" e_{FF}.

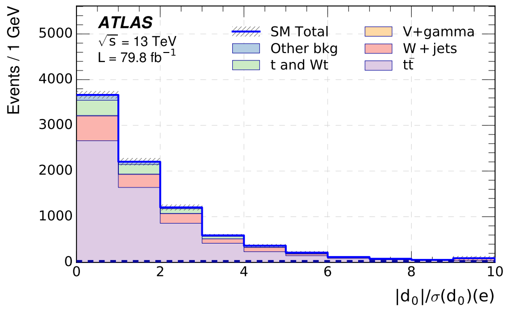


(f) Electron MR with $N_{\text{b-jets}} \geq 1$.
Truth matching requirement: "Bkg lep" e_{FF}.

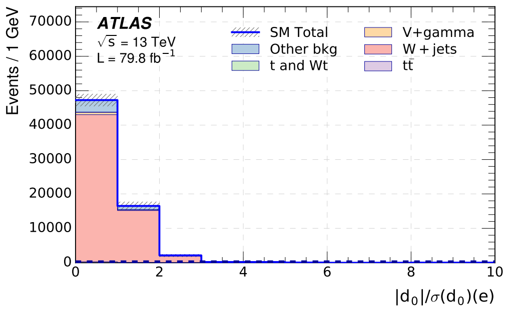
Figure 96: Distribution of the pseudorapidity: $|\eta|(e_{\text{FF}})$. All electron measurement region (MR) requirements are applied except for the requirement that e_{FF} satisfies the `Gradient` isolation selection, see Tab. 15. Truth matching is applied for the fake factor electron in each event and required to be: "non-isolated", "hadron", or "bkg lep". The blue line indicates the sum of the SM background processes and the dashed area the corresponding statistical uncertainty. The last bin in the histogram includes all events with values larger than the range of the histogram.



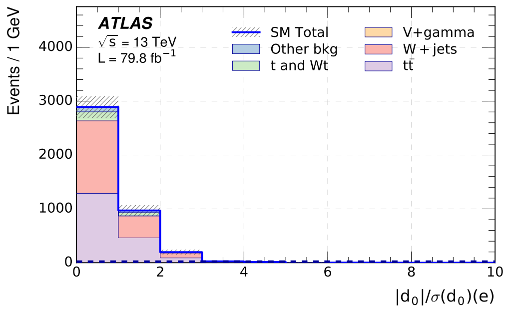
(a) Electron MR with $N_{\text{b-jets}} = 0$.
Truth matching requirement: "Non-isolated" e_{FF}.



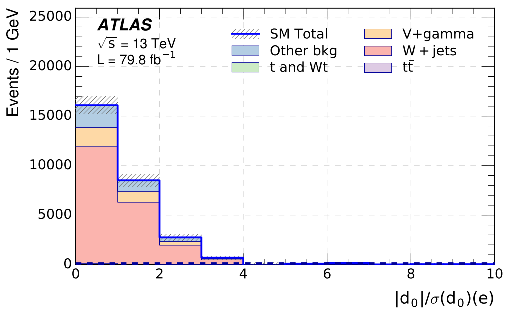
(b) Electron MR with $N_{\text{b-jets}} \geq 1$.
Truth matching requirement: "Non-isolated" e_{FF}.



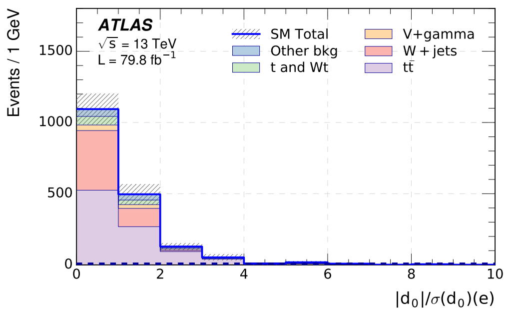
(c) Electron MR with $N_{\text{b-jets}} = 0$.
Truth matching requirement: "Hadron" e_{FF}.



(d) Electron MR with $N_{\text{b-jets}} \geq 1$.
Truth matching requirement: "Hadron" e_{FF}.

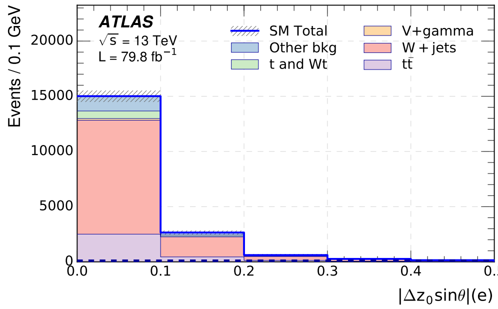


(e) Electron MR with $N_{\text{b-jets}} = 0$.
Truth matching requirement: "Bkg lep" e_{FF}.

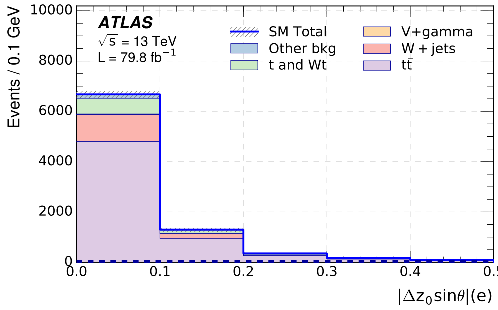


(f) Electron MR with $N_{\text{b-jets}} \geq 1$.
Truth matching requirement: "Bkg lep" e_{FF}.

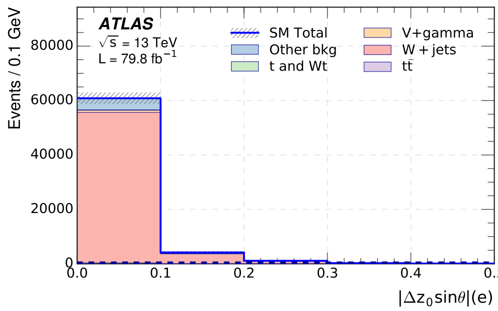
Figure 97: Distribution of the impact parameter: $|d_0|/\sigma(d_0)(e_{\text{FF}})$. All electron measurement region (MR) requirements are applied except for the requirement that e_{FF} satisfies the `Gradient` isolation selection, see Tab. 15. Truth matching is applied for the fake factor electron in each event and required to be: "non-isolated", "hadron", or "bkg lep". The blue line indicates the sum of the SM background processes and the dashed area the corresponding statistical uncertainty. The last bin in the histogram includes all events with values larger than the range of the histogram.



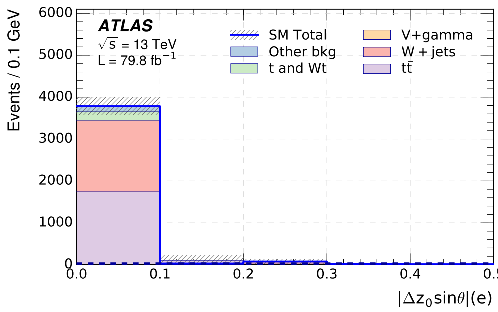
(a) Electron MR with $N_{\text{b-jets}} = 0$.
Truth matching requirement: "Non-isolated" e_{FF}.



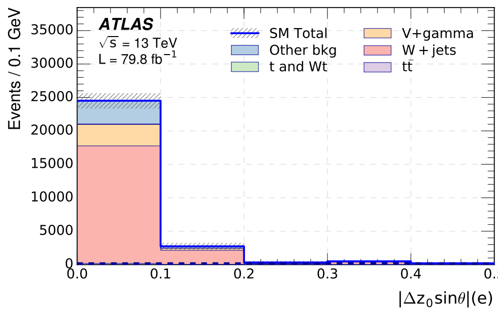
(b) Electron MR with $N_{\text{b-jets}} \geq 1$.
Truth matching requirement: "Non-isolated" e_{FF}.



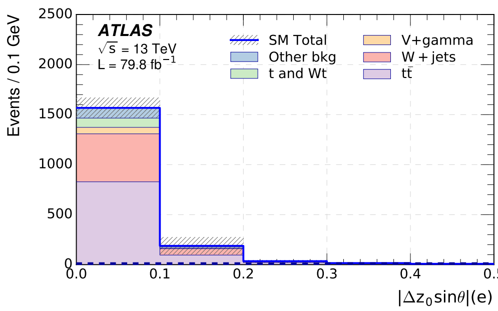
(c) Electron MR with $N_{\text{b-jets}} = 0$.
Truth matching requirement: "Hadron" e_{FF}.



(d) Electron MR with $N_{\text{b-jets}} \geq 1$.
Truth matching requirement: "Hadron" e_{FF}.

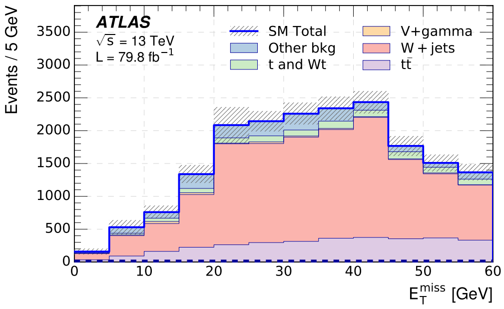


(e) Electron MR with $N_{\text{b-jets}} = 0$.
Truth matching requirement: "Bkg lep" e_{FF}.

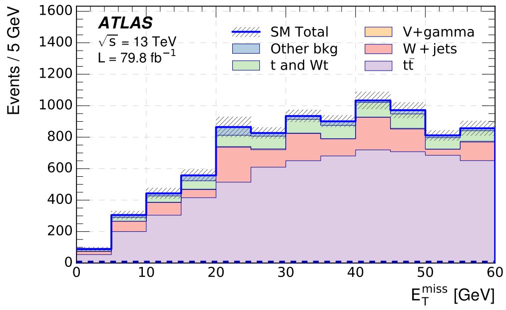


(f) Electron MR with $N_{\text{b-jets}} \geq 1$.
Truth matching requirement: "Bkg lep" e_{FF}.

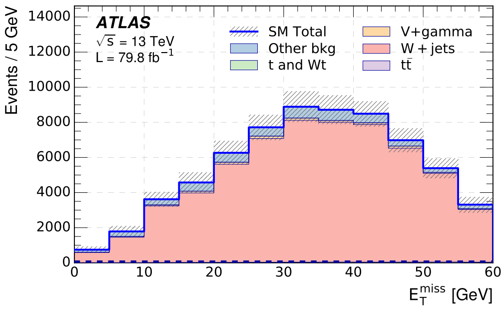
Figure 98: Distribution of the impact parameter variable: $|\Delta z_0 \sin(\theta)|(e_{\text{FF}})$. All electron MR requirements are applied except for the requirement that e_{FF} satisfies the `Gradient` isolation selection, see Tab. 15. Truth matching is applied for the fake factor electron in each event and required to be: "non-isolated", "hadron", or "bkg lep". The blue line indicates the sum of the SM background processes and the dashed area the corresponding statistical uncertainty. The last bin in the histogram includes all events with values larger than the range of the histogram.



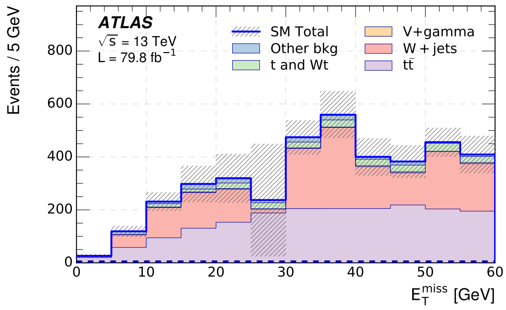
(a) Electron MR with $N_{\text{b-jets}} = 0$.
Truth matching requirement: "Non-isolated" e_{FF}.



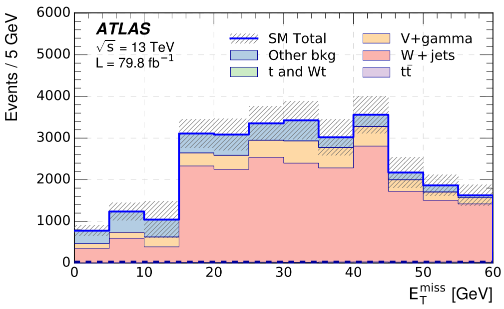
(b) Electron MR with $N_{\text{b-jets}} \geq 1$.
Truth matching requirement: "Non-isolated" e_{FF}.



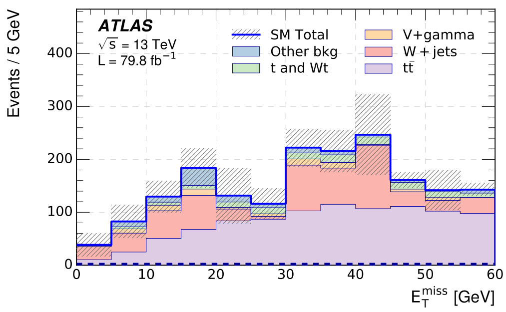
(c) Electron MR with $N_{\text{b-jets}} = 0$.
Truth matching requirement: "Hadron" e_{FF}.



(d) Electron MR with $N_{\text{b-jets}} \geq 1$.
Truth matching requirement: "Hadron" e_{FF}.

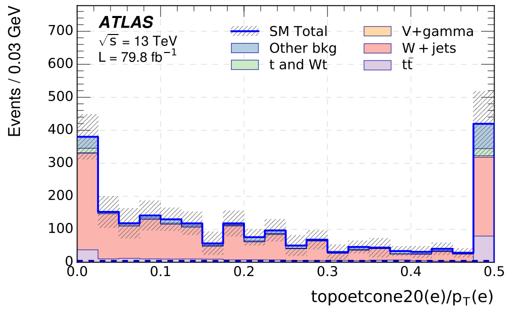


(e) Electron MR with $N_{\text{b-jets}} = 0$.
Truth matching requirement: "Bkg lep" e_{FF}.

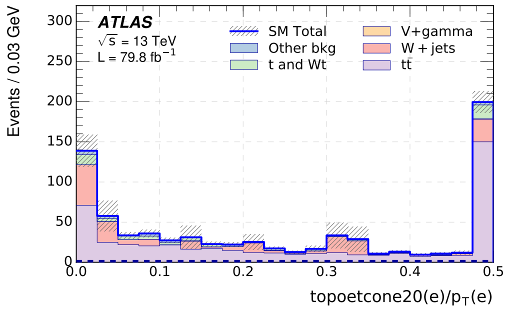


(f) Electron MR with $N_{\text{b-jets}} \geq 1$.
Truth matching requirement: "Bkg lep" e_{FF}.

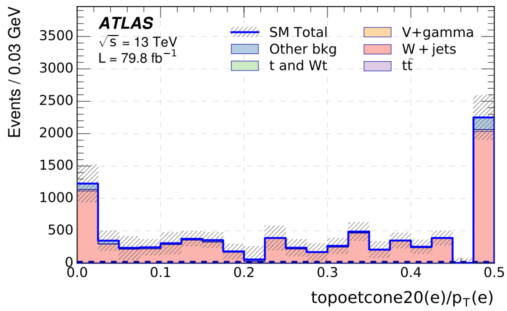
Figure 99: Distribution of the missing transverse energy: E_T^{miss}. All electron measurement region (MR) requirements are applied except for the requirement that e_{FF} satisfies the `Gradient` isolation selection, see Tab. 15. Additional requirements are used for e_{FF}: 10 GeV $< p_T <$ 15 GeV; $|d_0|/\sigma(d_0) < 1$; $|\Delta z_0 \sin(\theta)| < 0.1$ mm; $|\eta|$¡1. Truth matching is applied for the fake factor electron in each event and required to be: "non-isolated", "hadron", or "bkg lep". The blue line indicates the sum of the SM background processes and the dashed area the corresponding statistical uncertainty. The last bin in the histogram includes all events with values larger than the range of the histogram.



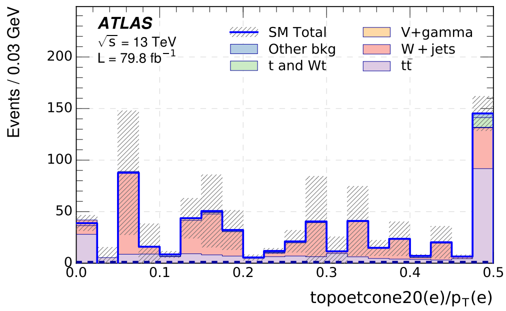
(a) Electron MR with $N_{\text{b-jets}} = 0$.
Truth matching requirement: "Non-isolated" e_{FF}.



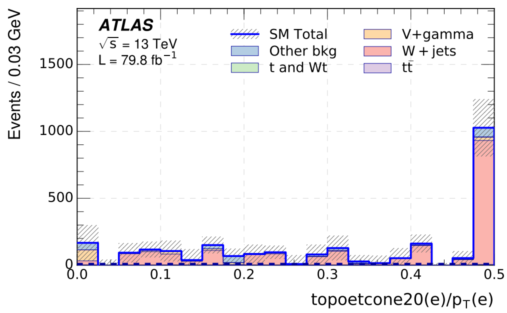
(b) Electron MR with $N_{\text{b-jets}} \geq 1$.
Truth matching requirement: "Non-isolated" e_{FF}.



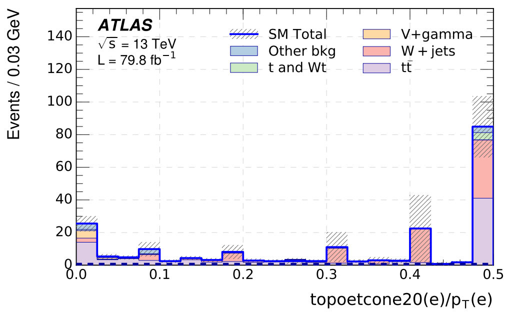
(c) Electron MR with $N_{\text{b-jets}} = 0$.
Truth matching requirement: "Hadron" e_{FF}.



(d) Electron MR with $N_{\text{b-jets}} \geq 1$.
Truth matching requirement: "Hadron" e_{FF}.

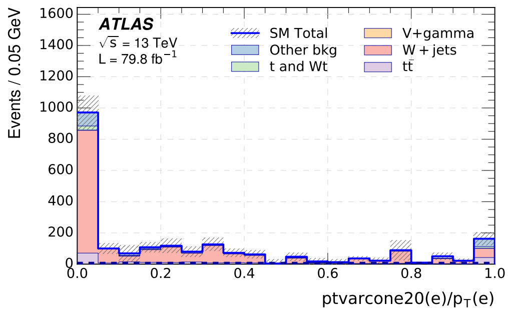


(e) Electron MR with $N_{\text{b-jets}} = 0$.
Truth matching requirement: "Bkg lep" e_{FF}.

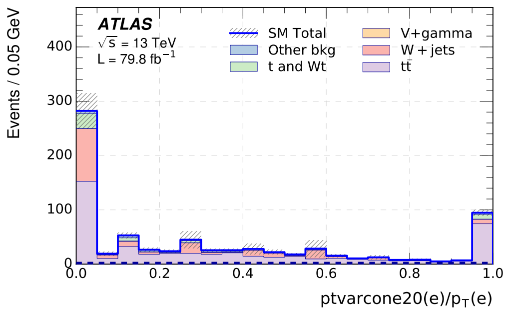


(f) Electron MR with $N_{\text{b-jets}} \geq 1$.
Truth matching requirement: "Bkg lep" e_{FF}.

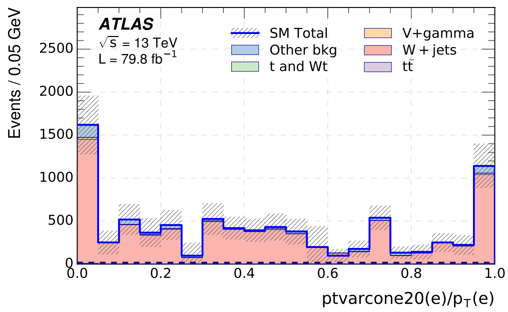
Figure 100: Distribution of the isolation variable: topoetcone20/$p_T(e_{\text{FF}})$. All electron measurement region (MR) requirements are applied except for the requirement that e_{FF} satisfies the `Gradient` isolation selection, see Tab. 15. Truth matching is applied for the fake factor electron in each event and required to be: "non-isolated", "hadron", or "bkg lep". The blue line indicates the sum of the SM background processes and the dashed area the corresponding statistical uncertainty. The last bin in the histogram includes all events with values larger than the range of the histogram.



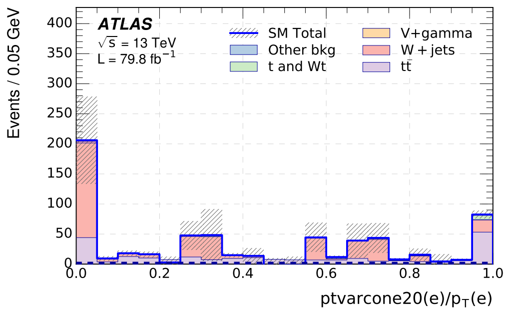
(a) Electron MR with $N_{\text{b-jets}} = 0$.
Truth matching requirement: "Non-isolated" e_{FF}.



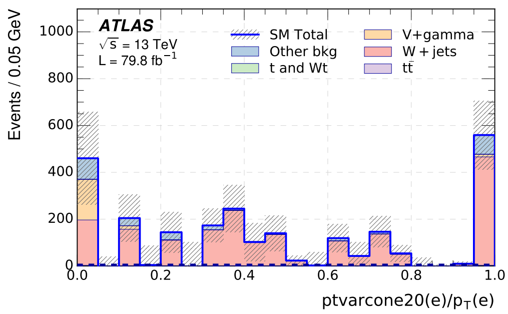
(b) Electron MR with $N_{\text{b-jets}} \geq 1$.
Truth matching requirement: "Non-isolated" e_{FF}.



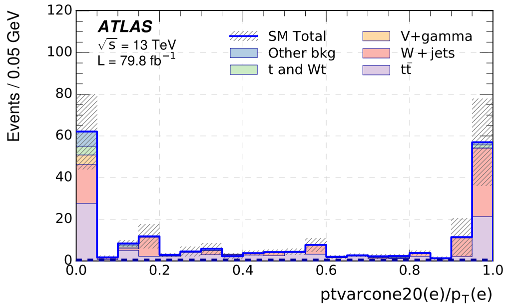
(c) Electron MR with $N_{\text{b-jets}} = 0$.
Truth matching requirement: "Hadron" e_{FF}.



(d) Electron MR with $N_{\text{b-jets}} \geq 1$.
Truth matching requirement: "Hadron" e_{FF}.



(e) Electron MR with $N_{\text{b-jets}} = 0$.
Truth matching requirement: "Bkg lep" e_{FF}.



(f) Electron MR with $N_{\text{b-jets}} \geq 1$.
Truth matching requirement: "Bkg lep" e_{FF}.

Figure 101: Distribution of the isolation variable: ptvarcone20/$p_T(e_{\text{FF}})$. All electron measurement region (MR) requirements are applied except for the requirement that e_{FF} satisfies the `Gradient` isolation selection, see Tab. 15. Additional requirements are used for e_{FF}: 10 GeV $< p_T <$ 15 GeV; $|d_0|/\sigma(d_0) < 1$; $|\Delta z_0 \sin(\theta)| < 0.1$ mm; $|\eta|$¡1. Truth matching is applied for the fake factor electron in each event and required to be: "non-isolated", "hadron", or "bkg lep". The blue line indicates the sum of the SM background processes and the dashed area the corresponding statistical uncertainty. The last bin in the histogram includes all events with values larger than the range of the histogram.

F Fake Factors

In this section, the studies applied to define the muon and electron fake factors are presented. The included figures and the final binning applied for each measurement region are summarised in tables below.

Table 70-72: Muon fake factors in the measurement regions applying `FixedCutLoose`, `Loosened FixedCutLoose` and `Loose` isolation respectively for μ_{FF}.
Table 73-74: Electron fake factors in the measurement regions applying `Gradient` and `Loose` isolation respectively for e_{FF}.

F.1 Muon Fake Factors: `FixedCutLoose` isolation

Fig.	$N_{\textbf{b-jets}}$	**Distribution**	**Binning**
102	0	$p_T(\mu_{\mathrm{FF}})$	No binning.
103	0	E_T^{miss}	Binning in $p_T(\mu_{\mathrm{FF}})$ applied.
104	0	$\lvert d_0\rvert/\sigma(d_0)(\mu_{\mathrm{FF}})$	Binning in $p_T(\mu_{\mathrm{FF}})$ applied.
105	0	N_{jets}	Binning in $p_T(\mu_{\mathrm{FF}})$ applied.
106	0	$\lvert d_0\rvert/\sigma(d_0)(\mu_{\mathrm{FF}})$	Binning in $p_T(\mu_{\mathrm{FF}})$ and $N_{\mathrm{jets}} \leq 2$. Final binning shown.
107	0	$\lvert d_0\rvert/\sigma(d_0)(\mu_{\mathrm{FF}})$	Binning in $p_T(\mu_{\mathrm{FF}})$ and $N_{\mathrm{jets}} > 2$. Final binning shown.
108	≥ 1	$p_T(\mu_{\mathrm{FF}})$, $\lvert\eta\rvert(\mu_{\mathrm{FF}})$, $\lvert d_0\rvert/\sigma(d_0)(\mu_{\mathrm{FF}})$, $\lvert\Delta z_0 \sin(\theta)\rvert(\mu_{\mathrm{FF}})$, N_{jets}, E_T^{miss}	No binning.
109	≥ 1	N_{jets}	Binning in $p_T(\mu_{\mathrm{FF}})$. Final bining shown.
		Final binning	
	0	Bins in $p_T(\mu_{\mathrm{FF}})$:	10 GeV $< p_T(\mu_{\mathrm{FF}}) \leq$ 20 GeV
	0		20 GeV $< p_T(\mu_{\mathrm{FF}}) \leq$ 30 GeV
	0		30 GeV $< p_T(\mu_{\mathrm{FF}}) \leq$ 40 GeV
	0		40 GeV $< p_T(\mu_{\mathrm{FF}}) \leq$ 50 GeV
	0	Bins in N_{jets}:	$N_{\mathrm{jets}} \leq 2$
	0		$N_{\mathrm{jets}} > 3$
	0	Bins in $\lvert d_0\rvert/\sigma(d_0)(\mu_{\mathrm{FF}})$:	$\lvert d_0\rvert/\sigma(d_0) < 2$
	0		$2 \leq \lvert d_0\rvert/\sigma(d_0) < 10$
	≥ 1	Bins in $p_T(\mu_{\mathrm{FF}})$:	10 GeV $< p_T(\mu_{\mathrm{FF}}) \leq$ 13 GeV
	≥ 1		13 GeV $< p_T(\mu_{\mathrm{FF}}) \leq$ 30 GeV
	≥ 1		30 GeV $< p_T(\mu_{\mathrm{FF}}) \leq$ 50 GeV
	≥ 1	Bins in N_{jets}:	$N_{\mathrm{jets}} \leq 2$
	≥ 1		$N_{\mathrm{jets}} > 3$

Table 70: Overview of the figures shown in the study of the muon fake factors using the measurement region with `FixedCutLoose` isolation for μ_{FF}. The final binning of the muon fake factors is listed.

F.2 Muon Fake Factors: `Loosened FixedCutLoose` isolation

Fig.	$N_{\text{b-jets}}$	Distribution	Binning
110	0	$p_T(\mu_{\rm FF})$, $\lvert\eta\rvert(\mu_{\rm FF})$, $\lvert d_0\rvert/\sigma(d_0)(\mu_{\rm FF})$, $\lvert\Delta z_0 \sin(\theta)\rvert(\mu_{\rm FF})$, $N_{\rm jets}$, $E_T^{\rm miss}$	No binning.
111	0	$\lvert d_0\rvert/\sigma(d_0)(\mu_{\rm FF})$	Binning in $p_T(\mu_{\rm FF})$ and $N_{\rm jets} \leq 2$. Final binning shown.
112	0	$\lvert d_0\rvert/\sigma(d_0)(\mu_{\rm FF})$	Binning in $p_T(\mu_{\rm FF})$ and $N_{\rm jets} > 2$ Final binning shown.
113	≥ 1	$p_T(\mu_{\rm FF})$, $\lvert\eta\rvert(\mu_{\rm FF})$, $\lvert d_0\rvert/\sigma(d_0)(\mu_{\rm FF})$, $\lvert\Delta z_0 \sin(\theta)\rvert(\mu_{\rm FF})$, $N_{\rm jets}$, $E_T^{\rm miss}$	No binning.
114	≥ 1	$N_{\rm jets}$	Binning in $p_T(\mu_{\rm FF})$. Final bining shown.
		Final binning	
	0	Bins in $p_T(\mu_{\rm FF})$:	10 GeV $< p_T(\mu_{\rm FF}) \leq$ 20 GeV
	0		20 GeV $< p_T(\mu_{\rm FF}) \leq$ 30 GeV
	0		30 GeV $< p_T(\mu_{\rm FF}) \leq$ 40 GeV
	0		40 GeV $< p_T(\mu_{\rm FF}) \leq$ 50 GeV
	0	Bins in $N_{\rm jets}$:	$N_{\rm jets} \leq 2$
	0		$N_{\rm jets} > 3$
	0	Bins in $\lvert d_0\rvert/\sigma(d_0)(\mu_{\rm FF})$:	$\lvert d_0\rvert/\sigma(d_0) < 2$
	0		$2 \leq \lvert d_0\rvert/\sigma(d_0) < 10$
	≥ 1	Bins in $p_T(\mu_{\rm FF})$:	10 GeV $< p_T(\mu_{\rm FF}) \leq$ 12 GeV
	≥ 1		12 GeV $< p_T(\mu_{\rm FF}) \leq$ 30 GeV
	≥ 1		30 GeV $< p_T(\mu_{\rm FF}) \leq$ 50 GeV
	≥ 1	Bins in $N_{\rm jets}$:	$N_{\rm jets} \leq 2$
	≥ 1		$N_{\rm jets} > 3$

Table 71: Overview of the figures shown in the study of the muon fake factors using the measurement region with `Loosened FixedCutLoose` isolation for $\mu_{\rm FF}$. The final binning of the muon fake factors is listed.

F.3 Muon Fake Factors: `Loose` isolation

Fig.	$N_{\textbf{b-jets}}$	Distribution	Binning
115	0	$p_T(\mu_{\mathrm{FF}})$, $\lvert\eta\rvert(\mu_{\mathrm{FF}})$, $\lvert d_0\rvert/\sigma(d_0)(\mu_{\mathrm{FF}})$, $\lvert\Delta z_0 \sin(\theta)\rvert(\mu_{\mathrm{FF}})$, N_{jets}, E_T^{miss}	No binning.
116	0	$\lvert\eta\rvert(\mu_{\mathrm{FF}})$	Binning in $p_T(\mu_{\mathrm{FF}})$ and $N_{\mathrm{jets}} \leq 2$. Final binning shown.
117	0	$\lvert\eta\rvert(\mu_{\mathrm{FF}})$	Binning in $p_T(\mu_{\mathrm{FF}})$ and $N_{\mathrm{jets}} > 2$. Final binning shown.
118	≥ 1	$p_T(\mu_{\mathrm{FF}})$, $\lvert\eta\rvert(\mu_{\mathrm{FF}})$, $\lvert d_0\rvert/\sigma(d_0)(\mu_{\mathrm{FF}})$, $\lvert\Delta z_0 \sin(\theta)\rvert(\mu_{\mathrm{FF}})$, N_{jets}, E_T^{miss}	No binning.
119	≥ 1	$\lvert\eta\rvert(\mu_{\mathrm{FF}})$	Binning in $p_T(\mu_{\mathrm{FF}})$. Final bining shown.
		Final binning	
	0	Bins in $p_T(\mu_{\mathrm{FF}})$:	10 GeV $< p_T(\mu_{\mathrm{FF}}) \leq$ 12 GeV
	0		12 GeV $< p_T(\mu_{\mathrm{FF}}) \leq$ 15 GeV
	0		15 GeV $< p_T(\mu_{\mathrm{FF}}) \leq$ 20 GeV
	0		20 GeV $< p_T(\mu_{\mathrm{FF}}) \leq$ 26 GeV
	0		26 GeV $< p_T(\mu_{\mathrm{FF}}) \leq$ 50 GeV
	0	Bins in N_{jets}:	$N_{\mathrm{jets}} \leq 2$
	0		$N_{\mathrm{jets}} > 3$
	0	Bins in $\lvert\eta\rvert(\mu_{\mathrm{FF}})$:	$\lvert\eta\rvert < 0.75$
	0		$0.75 \leq \lvert\eta\rvert(\mu_{\mathrm{FF}}) < 1.25$
	0		$1.25 \leq \lvert\eta\rvert(\mu_{\mathrm{FF}}) < 2.5$
	0	Bins in $p_T(\mu_{\mathrm{FF}})$:	10 GeV $< p_T(\mu_{\mathrm{FF}}) \leq$ 12 GeV
	≥ 1		12 GeV $< p_T(\mu_{\mathrm{FF}}) \leq$ 15 GeV
	≥ 1		15 GeV $< p_T(\mu_{\mathrm{FF}}) \leq$ 20 GeV
	≥ 1		20 GeV $< p_T(\mu_{\mathrm{FF}}) \leq$ 26 GeV
	≥ 1		26 GeV $< p_T(\mu_{\mathrm{FF}}) \leq$ 50 GeV
	≥ 1	Bins in $\lvert\eta\rvert(\mu_{\mathrm{FF}})$:	$\lvert\eta\rvert < 0.75$
	≥ 1		$0.75 \leq \lvert\eta\rvert(\mu_{\mathrm{FF}}) < 1.25$
	≥ 1		$1.25 \leq \lvert\eta\rvert(\mu_{\mathrm{FF}}) < 2.5$

Table 72: Overview of the figures shown in the study of the muon fake factors using the measurement region with `Loose` isolation for μ_{FF}. The final binning of the muon fake factors is listed.

F.4 Elecctron Fake Factors: `Gradient` isolation

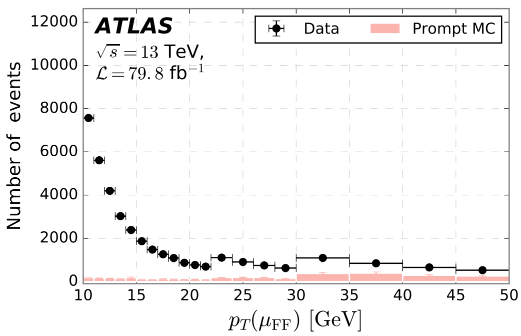
Fig.	$N_{\text{b-jets}}$	Distribution	Binning
121	0	$p_T(e_{\text{FF}})$	No binning.
122	0	$\lvert d_0\rvert/\sigma(d_0)(e_{\text{FF}})$, $\lvert\Delta z_0 \sin(\theta)\rvert(e_{\text{FF}})$	Binning in $p_T(e_{\text{FF}})$ applied.
123	0	$\lvert\eta\rvert(e_{\text{FF}})$, E_T^{miss}	Binning in $p_T(e_{\text{FF}})$ applied.
124	0	N_{jets}	Binning in $p_T(e_{\text{FF}})$ applied.
125	0	$\lvert d_0\rvert/\sigma(d_0)(e_{\text{FF}})$	Binning in $p_T(e_{\text{FF}})$ and $N_{\text{jets}} \leq 2$. Final binning shown.
126	0	$\lvert d_0\rvert/\sigma(d_0)(e_{\text{FF}})$	Binning in $p_T(e_{\text{FF}})$ and $N_{\text{jets}} > 2$. Final binning shown.
127	≥ 1	$p_T(e_{\text{FF}})$, $\lvert\eta\rvert(\mu_{\text{FF}})$, $\lvert d_0\rvert/\sigma(d_0)(e_{\text{FF}})$, $\lvert\Delta z_0 \sin(\theta)\rvert(e_{\text{FF}})$, N_{jets}, E_T^{miss}	No binning.
128	≥ 1	$\lvert d_0\rvert/\sigma(d_0)(e_{\text{FF}})$	Binning in $p_T(e_{\text{FF}})$. Final bining shown.
		Final binning	
	0	Bins in $p_T(e_{\text{FF}})$:	10 GeV $< p_T(e_{\text{FF}}) \leq 13$ GeV
	0		13 GeV $< p_T(e_{\text{FF}}) \leq 15$ GeV
	0		15 GeV $< p_T(e_{\text{FF}}) \leq 17$ GeV
	0		17 GeV $< p_T(e_{\text{FF}}) \leq 35$ GeV
	0		35 GeV $< p_T(e_{\text{FF}}) \leq 50$ GeV
	0	Bins in N_{jets}:	$N_{\text{jets}} \leq 2$
	0		$N_{\text{jets}} > 3$
	0	Bins in $\lvert d_0\rvert/\sigma(d_0)(e_{\text{FF}})$:	$\lvert d_0\rvert/\sigma(d_0)(e_{\text{FF}}) < 3$
	0		$3 \leq \lvert d_0\rvert/\sigma(d_0)(e_{\text{FF}}) < 10$
	≥ 1	Bins in $p_T(e_{\text{FF}})$:	10 GeV $< p_T(e_{\text{FF}}) \leq 15$ GeV
	≥ 1		15 GeV $< p_T(e_{\text{FF}}) \leq 25$ GeV
	≥ 1		25 GeV $< p_T(e_{\text{FF}}) \leq 50$ GeV
	≥ 1	Bins in $\lvert d_0\rvert/\sigma(d_0)(e_{\text{FF}})$:	$\lvert d_0\rvert/\sigma(d_0)(e_{\text{FF}}) < 3$
	≥ 1		$3 \leq \lvert d_0\rvert/\sigma(d_0)(e_{\text{FF}}) < 10$

Table 73: Overview of the figures shown in the study of the electron fake factors using the measurement region with `Gradient` isolation for e_{FF}. The final binning of the electron fake factors is listed.

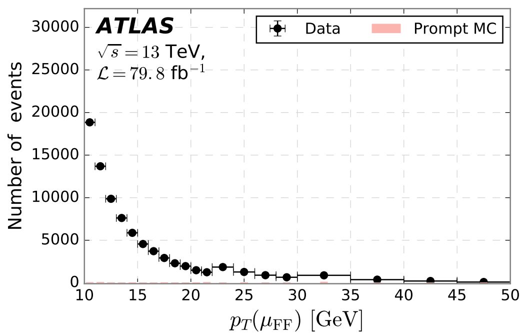
F.5 Electron Fake Factors: `Loose` isolation

Fig.	$N_{\text{b-jets}}$	Distribution	Binning
129	0	$p_T(\mu_{\mathrm{FF}})$, $\lvert\eta\rvert(\mu_{\mathrm{FF}})$, $\lvert d_0\rvert/\sigma(d_0)(\mu_{\mathrm{FF}})$, $\lvert\Delta z_0 \sin(\theta)\rvert(\mu_{\mathrm{FF}})$, N_{jets}, E_T^{miss}	No binning.
130	0	$\lvert d_0\rvert/\sigma(d_0)(e_{\mathrm{FF}})$	Binning in $p_T(e_{\mathrm{FF}})$ and $N_{\mathrm{jets}} \leq 2$. Final binning shown.
131	0	$\lvert d_0\rvert/\sigma(d_0)(e_{\mathrm{FF}})$	Binning in $p_T(e_{\mathrm{FF}})$ and $N_{\mathrm{jets}} > 2$. Final binning shown.
132	≥ 1	$p_T(\mu_{\mathrm{FF}})$, $\lvert\eta\rvert(\mu_{\mathrm{FF}})$, $\lvert d_0\rvert/\sigma(d_0)(\mu_{\mathrm{FF}})$, $\lvert\Delta z_0 \sin(\theta)\rvert(\mu_{\mathrm{FF}})$, N_{jets}, E_T^{miss}	No binning.
133	≥ 1	$\lvert d_0\rvert/\sigma(d_0)(e_{\mathrm{FF}})$	Binning in $p_T(e_{\mathrm{FF}})$ and $N_{\mathrm{jets}} \leq 2$. Final binning shown.
134	≥ 1	$\lvert d_0\rvert/\sigma(d_0)(e_{\mathrm{FF}})$	Binning in $p_T(e_{\mathrm{FF}})$ and $N_{\mathrm{jets}} > 2$. Final binning shown.
		Final binning	
	0	Bins in $p_T(e_{\mathrm{FF}})$:	10 GeV $< p_T(e_{\mathrm{FF}}) \leq$ 13 GeV
	0		13 GeV $< p_T(e_{\mathrm{FF}}) \leq$ 15 GeV
	0		15 GeV $< p_T(e_{\mathrm{FF}}) \leq$ 17 GeV
	0		20 GeV $< p_T(e_{\mathrm{FF}}) \leq$ 20 GeV
	0		20 GeV $< p_T(e_{\mathrm{FF}}) \leq$ 50 GeV
	0	Bins in N_{jets}:	$N_{\mathrm{jets}} \leq 2$
	0		$N_{\mathrm{jets}} > 3$
	0	Bins in $\lvert d_0\rvert/\sigma(d_0)(e_{\mathrm{FF}})$:	$\lvert d_0\rvert/\sigma(d_0)(e_{\mathrm{FF}}) < 3$
	0		$3 \leq \lvert d_0\rvert/\sigma(d_0)(e_{\mathrm{FF}}) < 10$
	≥ 1	Bins in $p_T(e_{\mathrm{FF}})$:	10 GeV $< p_T(e_{\mathrm{FF}}) \leq$ 15 GeV
	≥ 1		15 GeV $< p_T(e_{\mathrm{FF}}) \leq$ 22 GeV
	≥ 1		22 GeV $< p_T(e_{\mathrm{FF}}) \leq$ 50 GeV
	≥ 1	Bins in N_{jets}:	$N_{\mathrm{jets}} \leq 2$
	≥ 1		$N_{\mathrm{jets}} > 3$
	≥ 1	Bins in $\lvert d_0\rvert/\sigma(d_0)(e_{\mathrm{FF}})$:	$\lvert d_0\rvert/\sigma(d_0)(e_{\mathrm{FF}}) < 3$
	≥ 1		$3 \leq \lvert d_0\rvert/\sigma(d_0)(e_{\mathrm{FF}}) < 10$

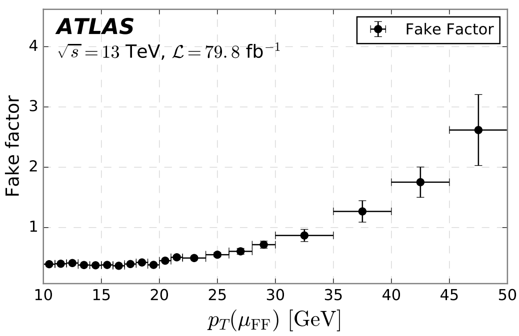
Table 74: Overview of the figures shown in the study of the electron fake factors using the measurement region with `Loose` isolation for e_{FF}. The final binning of the electron fake factors is listed.



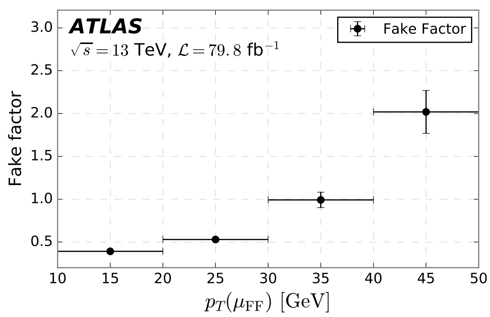
(a) Fake factor numerator: ID for $\mu_{\rm FF}$. Prompt MC: Sum of all MC background processes. $\mu_{\rm FF}$ is required to be real.



(b) Fake factor denominator: anti-ID for $\mu_{\rm FF}$. Prompt MC: Sum of all MC background processes. $\mu_{\rm FF}$ is required to be real.

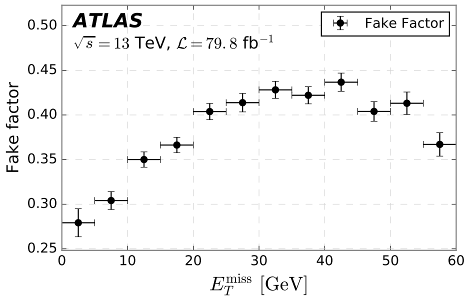


(c) Fake factor as a function of $p_T(\mu_{\rm FF})$.

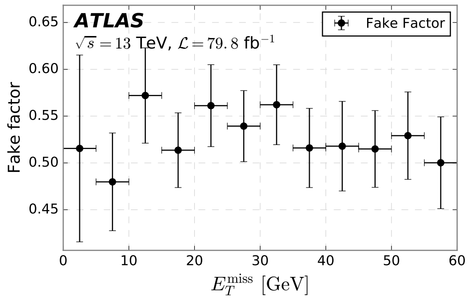


(d) Applied binning for Fig. 102c.

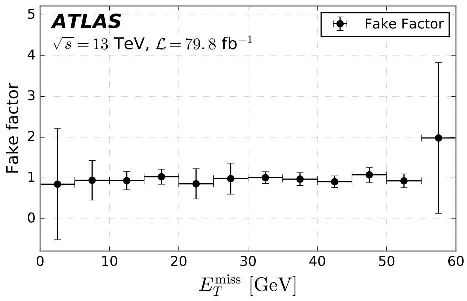
Figure 102: Muon fake factor as a function of $p_T(\mu_{\rm FF})$. The ID and anti-ID criteria are defined using the `FixedCutLoose` isolation in the muon measurement region with $N_{\text{b-jets}} = 0$.



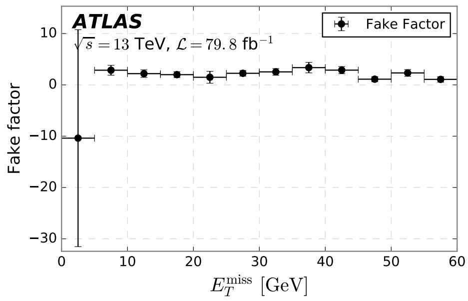
(a) Bin: 10 GeV $\leq p_T(\mu_{\text{FF}}) < 20$ GeV.



(b) Bin: 10 GeV $\leq p_T(\mu_{\text{FF}}) < 20$ GeV.

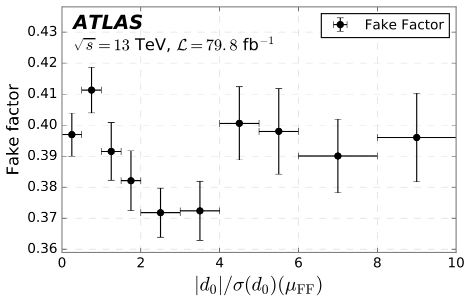


(c) Bin: 10 GeV $\leq p_T(\mu_{\text{FF}}) < 20$ GeV.

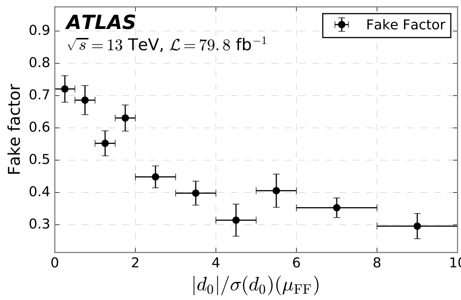


(d) Bin: 10 GeV $\leq p_T(\mu_{\text{FF}}) < 20$ GeV.

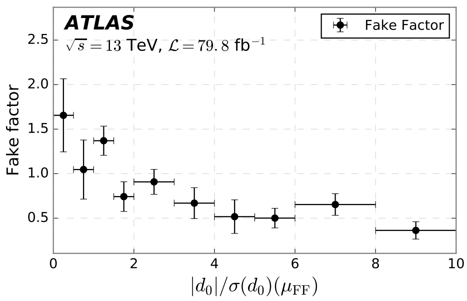
Figure 103: Muon fake factor as a function of E_T^{miss}. The considered fake factors are binned in $p_T(\mu_{\text{FF}})$. The ID and anti-ID criteria are defined using the `FixedCutLoose` isolation in the muon measurement region with $N_{\text{b-jets}} = 0$.



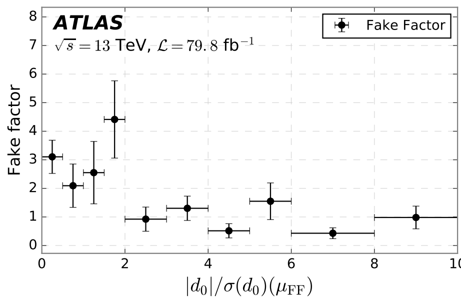
(a) Bin: 10 GeV $\leq p_T(\mu_{\text{FF}}) < 20$ GeV.



(b) Bin: 10 GeV $\leq p_T(\mu_{\text{FF}}) < 20$ GeV.

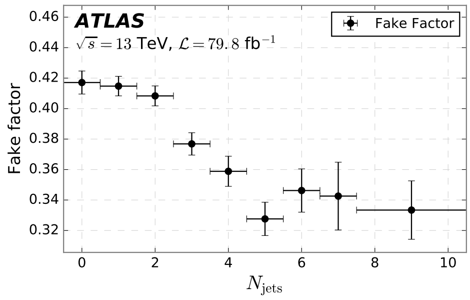


(c) Bin: 10 GeV $\leq p_T(\mu_{\text{FF}}) < 20$ GeV.

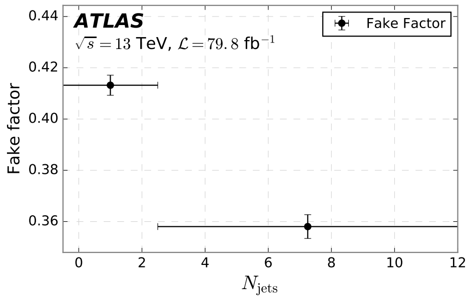


(d) Bin: 10 GeV $\leq p_T(\mu_{\text{FF}}) < 20$ GeV.

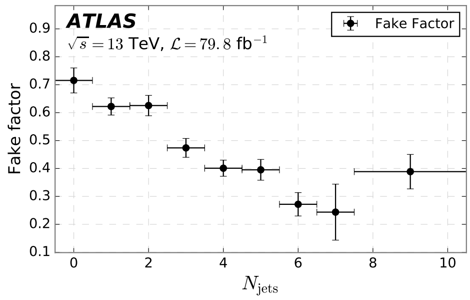
Figure 104: Muon fake factor as a function of $|d_0|/\sigma(d_0)(\mu_{\text{FF}})$. The considered fake factors are binned in $p_T(\mu_{\text{FF}})$. The ID and anti-ID criteria are defined using the `FixedCutLoose` isolation in the muon measurement region with $N_{\text{b-jets}} = 0$.



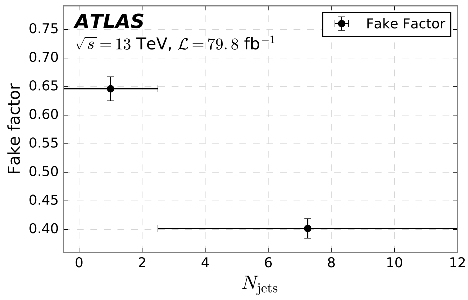
(a) Fake factor as a function of N_{jets}. Bin: 10 GeV $\leq p_T(\mu_{\text{FF}}) <$ 20 GeV.



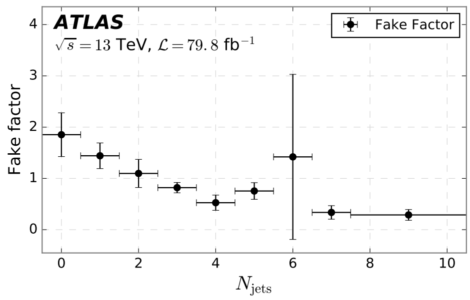
(b) Applied binning as a function of N_{jets} for Fig. 105a.



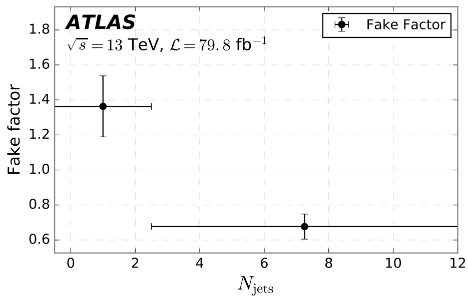
(c) Fake factor as a function of N_{jets}. Bin: 20 GeV $\leq p_T(\mu_{\text{FF}}) <$ 30 GeV.



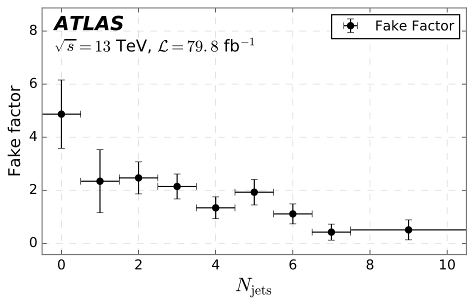
(d) Applied binning as a function of N_{jets} for Fig. 105c.



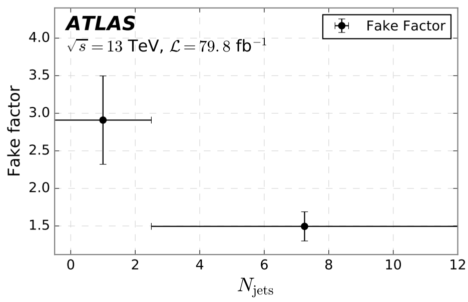
(e) Fake factor as a function of N_{jets}. Bin: 30 GeV $\leq p_T(\mu_{\text{FF}}) <$ 40 GeV.



(f) Applied binning as a function of N_{jets} for Fig. 105e.

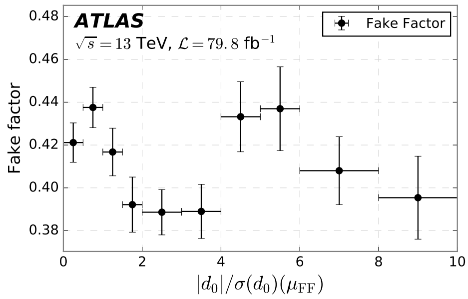


(g) Fake factor as a function of N_{jets}. Bin: 40 GeV $\leq p_T(\mu_{\text{FF}}) <$ 50 GeV.

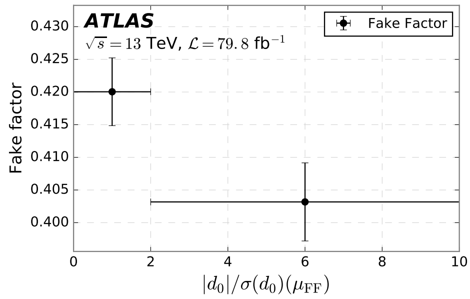


(h) Applied binning as a function of N_{jets} for Fig. 105g.

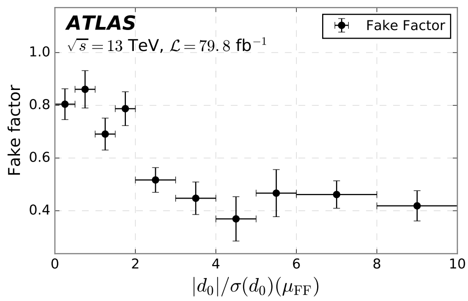
Figure 105: Muon fake factor as a function of N_{jets}. The considered fake factors are binned in $p_T(\mu_{\text{FF}})$. The ID and anti-ID are criteria defined using the `FixedCutLoose` isolation in the muon measurement region with $N_{\text{b-jets}} = 0$.



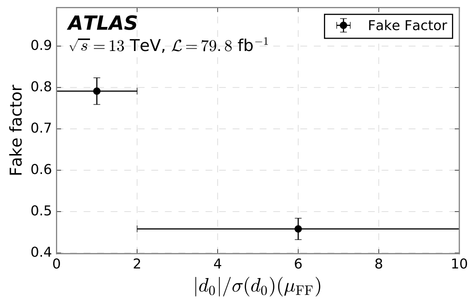
(a) Bin: 10 GeV $\leq p_T(\mu_{\mathrm{FF}}) <$ 20 GeV and $N_{\mathrm{jets}} \leq 2$.



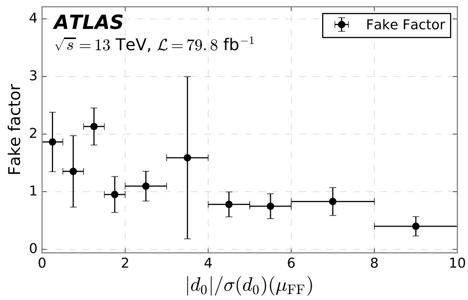
(b) Applied binning as a function of $|d_0|/\sigma(d_0)(\mu_{\mathrm{FF}})$ for Fig. 106a.



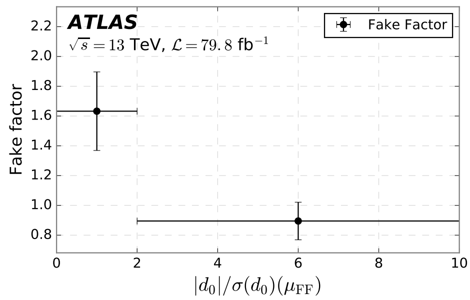
(c) Bin: 20 GeV $\leq p_T(\mu_{\mathrm{FF}}) <$ 30 GeV and $N_{\mathrm{jets}} \leq 2$.



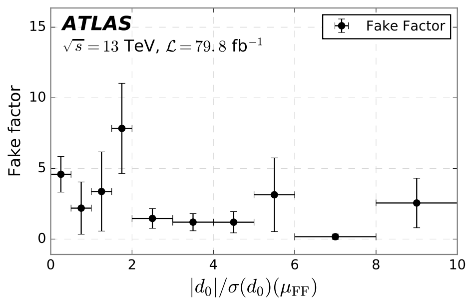
(d) Applied binning as a function of $|d_0|/\sigma(d_0)(\mu_{\mathrm{FF}})$ for Fig. 106c.



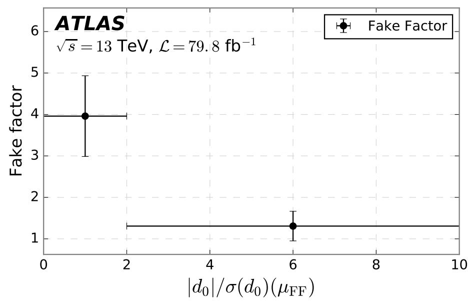
(e) Bin: 30 GeV $\leq p_T(\mu_{\mathrm{FF}}) <$ 40 GeV and $N_{\mathrm{jets}} \leq 2$.



(f) Applied binning as a function of $|d_0|/\sigma(d_0)(\mu_{\mathrm{FF}})$ for Fig. 106e.

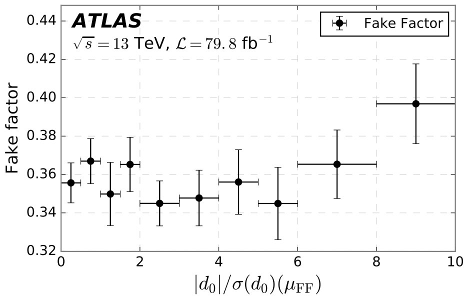


(g) Bin: 40 GeV $\leq p_T(\mu_{\mathrm{FF}}) <$ 50 GeV and $N_{\mathrm{jets}} \leq 2$.

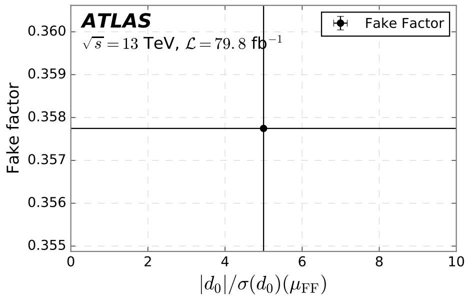


(h) Applied binning as a function of $|d_0|/\sigma(d_0)(\mu_{\mathrm{FF}})$ for Fig. 106g.

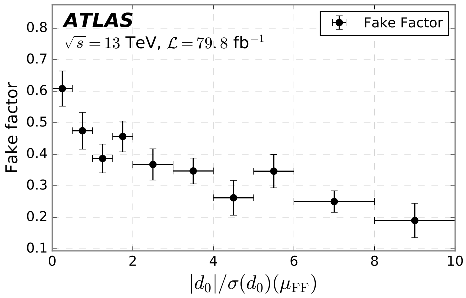
Figure 106: Muon fake factor as a function of $|d_0|/\sigma(d_0)(\mu_{\mathrm{FF}})$. The considered fake factors are binned in $p_T(\mu_{\mathrm{FF}})$ and N_{jets}. The ID and anti-ID criteria are defined using the `FixedCutLoose` isolation in the muon measurement region with $N_{\text{b-jets}} = 0$.



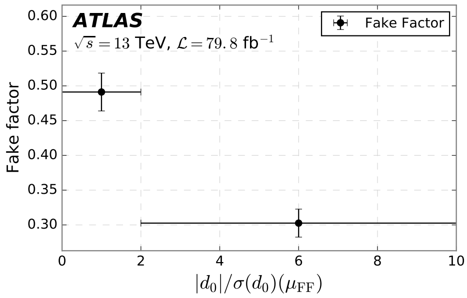
(a) Bin: 10 GeV $\leq p_T(\mu_{\mathrm{FF}}) <$ 20 GeV and $N_{\mathrm{jets}} > 2$.



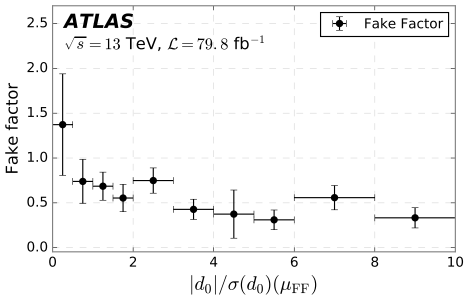
(b) Applied binning as a function of $|d_0|/\sigma(d_0)(\mu_{\mathrm{FF}})$ for Fig. 107a.



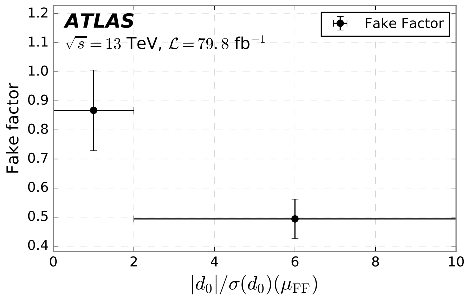
(c) Bin: 20 GeV $\leq p_T(\mu_{\mathrm{FF}}) <$ 30 GeV and $N_{\mathrm{jets}} > 2$.



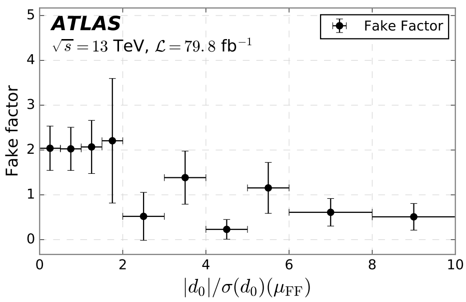
(d) Applied binning as a function of $|d_0|/\sigma(d_0)(\mu_{\mathrm{FF}})$ for Fig. 107c.



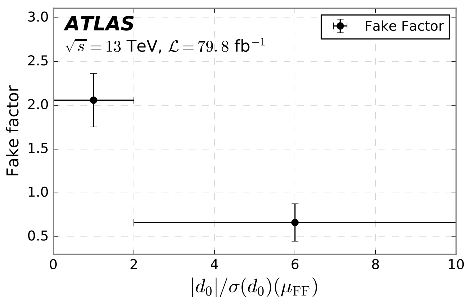
(e) Bin: 30 GeV $\leq p_T(\mu_{\mathrm{FF}}) <$ 40 GeV and $N_{\mathrm{jets}} > 2$.



(f) Applied binning as a function of $|d_0|/\sigma(d_0)(\mu_{\mathrm{FF}})$ for Fig. 107e.

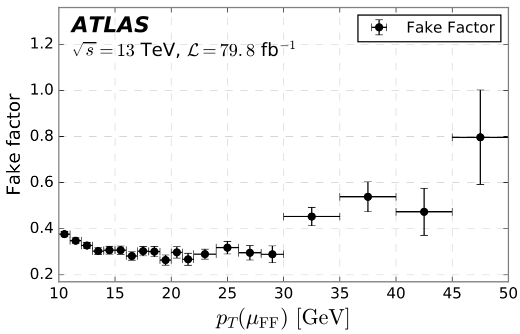


(g) Bin: 40 GeV $\leq p_T(\mu_{\mathrm{FF}}) <$ 50 GeV and $N_{\mathrm{jets}} > 2$.

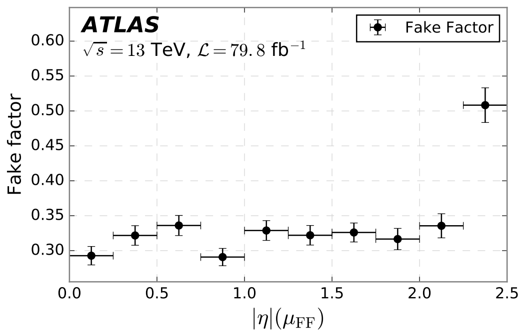


(h) Applied binning as a function of $|d_0|/\sigma(d_0)(\mu_{\mathrm{FF}})$ for Fig. 107g.

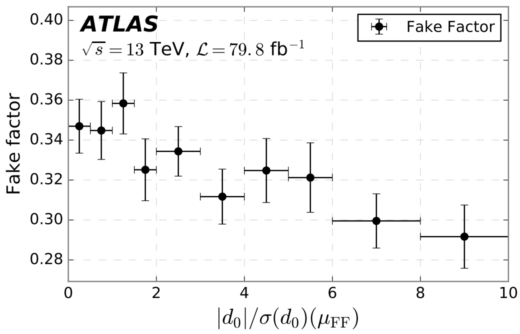
Figure 107: Muon fake factor as a function of $|d_0|/\sigma(d_0)(\mu_{\mathrm{FF}})$. The considered fake factors are binned in $p_T(\mu_{\mathrm{FF}})$ and N_{jets}. The ID and anti-ID criteria are defined using the `FixedCutLoose` isolation in the muon measurement region with $N_{\text{b-jets}} = 0$.



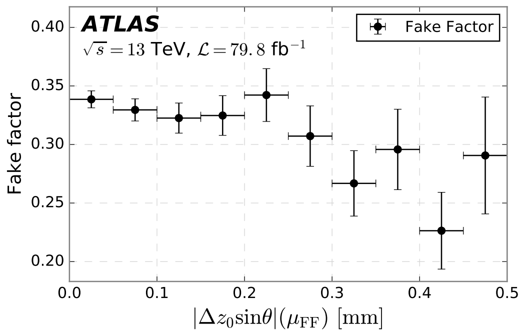
(a) Fake factor as a function of $p_T(\mu_{\mathrm{FF}})$.



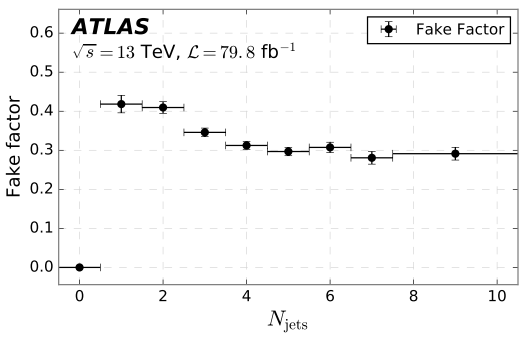
(b) Fake factor as a function of $|\eta|(\mu_{\mathrm{FF}})$.



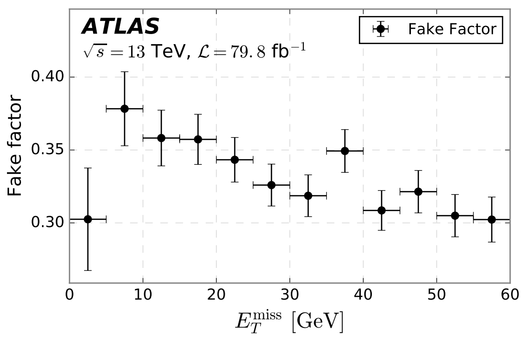
(c) Fake factor as a function of $|d_0|/\sigma(d_0)(\mu_{\mathrm{FF}})$.



(d) Fake factor as a function of $|\Delta z_0 \sin(\theta)|(\mu_{\mathrm{FF}})$.

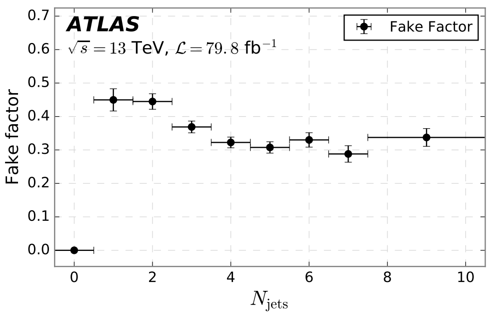


(e) Fake factor as a function of N_{jets}.

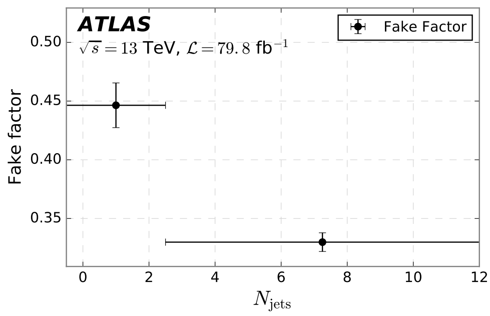


(f) Fake factor as a function of E_T^{miss}.

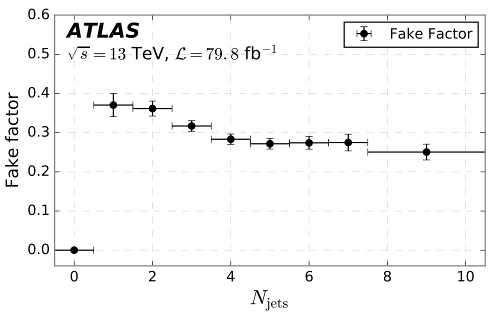
Figure 108: Muon fake factor as a function of: $p_T(\mu_{\mathrm{FF}})$; $|\eta|(\mu_{\mathrm{FF}})$; $|d_0|/\sigma(d_0)(\mu_{\mathrm{FF}})$; $|\Delta z_0 \sin(\theta)|(\mu_{\mathrm{FF}})$; N_{jets}; and E_T^{miss}. The ID and anti-ID criteria are defined using the `FixedCutLoose` isolation in the muon measurement region with $N_{\text{b-jets}} \geq 1$.



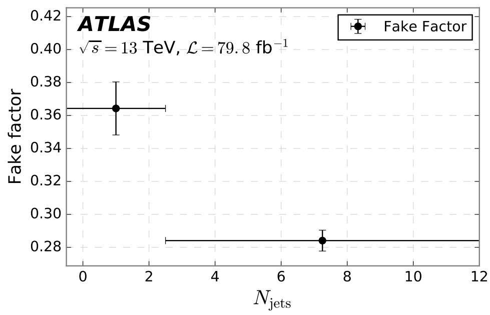
(a) Fake factor as a function of N_{jets}. Bin: 10 GeV $\leq p_T(\mu_{\mathrm{FF}}) <$ 20 GeV.



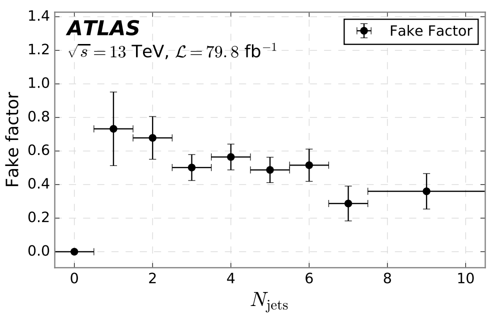
(b) Applied binning as a function of N_{jets} for Fig. 109a.



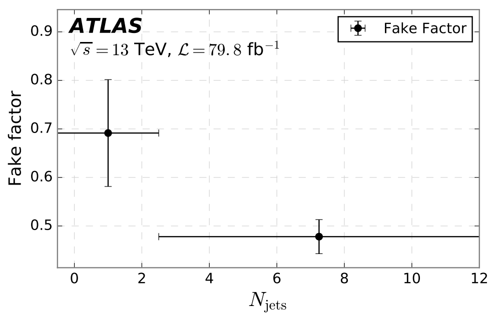
(c) Fake factor as a function of N_{jets}. Bin: 20 GeV $\leq p_T(\mu_{\mathrm{FF}}) <$ 30 GeV.



(d) Applied binning as a function of N_{jets} for Fig. 109c.

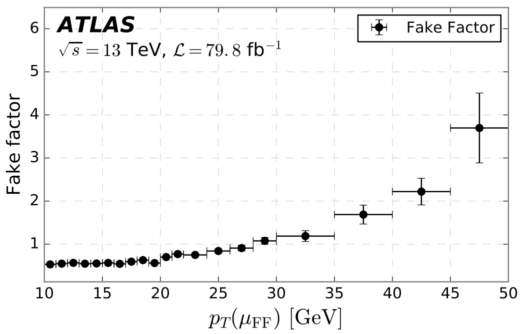


(e) Fake factor as a function of N_{jets}. Bin: 40 GeV $\leq p_T(\mu_{\mathrm{FF}}) <$ 50 GeV.

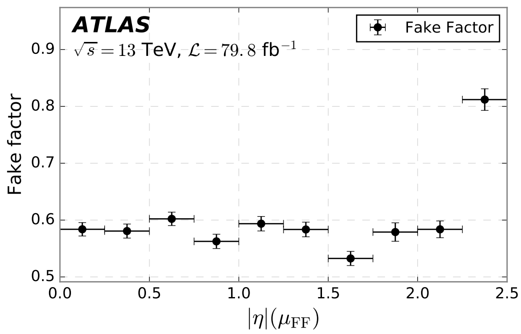


(f) Applied binning as a function of N_{jets} for Fig. 109e.

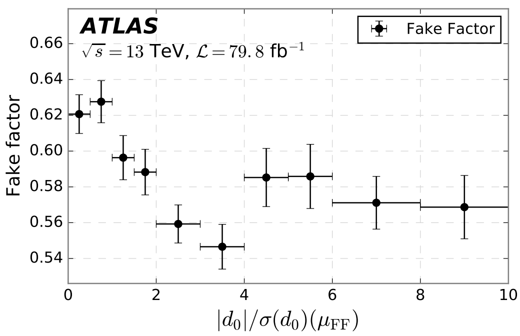
Figure 109: Muon fake factor as a function of N_{jets}. The considered fake factors are binned in $p_T(\mu_{\mathrm{FF}})$. The ID and anti-ID criteria are defined using the `FixedCutLoose` isolation selection in the muon measurement region with $N_{\text{b-jets}} \geq 1$.



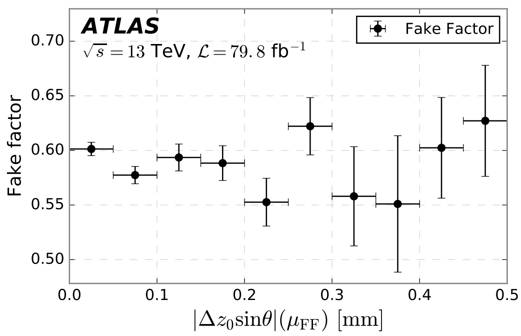
(a) Fake factor as a function of $p_T(\mu_{\mathrm{FF}})$.



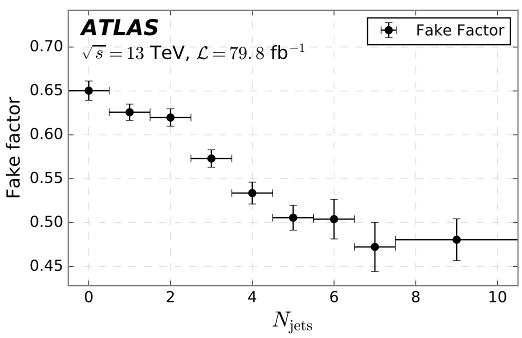
(b) Fake factor as a function of $|\eta|(\mu_{\mathrm{FF}})$.



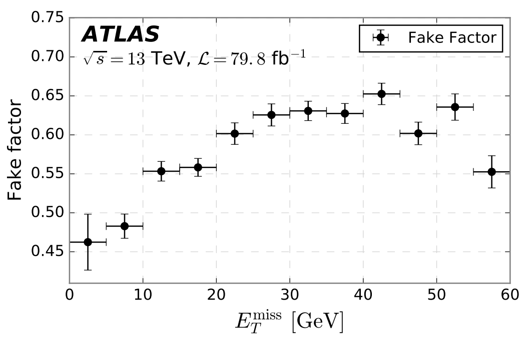
(c) Fake factor as a function of $|d_0|/\sigma(d_0)(\mu_{\mathrm{FF}})$.



(d) Fake factor as a function of $|\Delta z_0 \sin(\theta)|(\mu_{\mathrm{FF}})$.

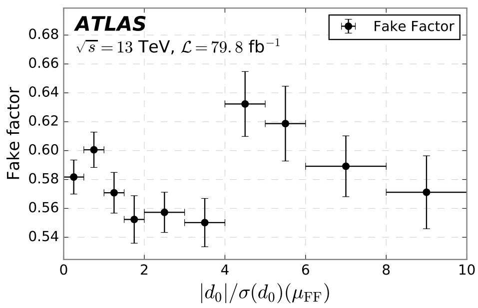


(e) Fake factor as a function of N_{jets}.

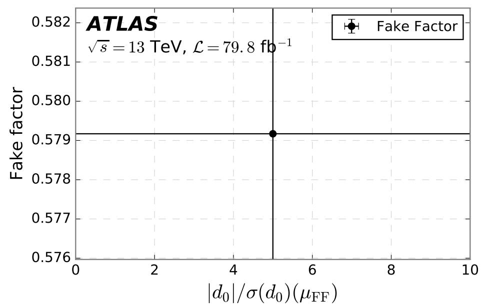


(f) Fake factor as a function of E_T^{miss}.

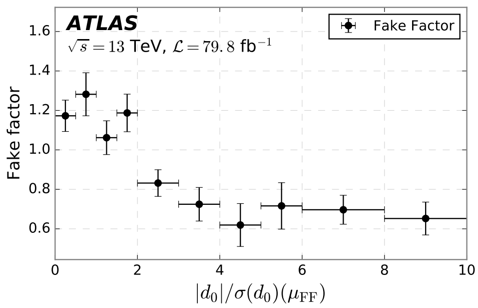
Figure 110: Muon fake factor as a function of: $p_T(\mu_{\mathrm{FF}})$; $|\eta|(\mu_{\mathrm{FF}})$; $|d_0|/\sigma(d_0)(\mu_{\mathrm{FF}})$; $|\Delta z_0 \sin(\theta)|(\mu_{\mathrm{FF}})$; N_{jets}; and E_T^{miss}. The ID and anti-ID criteria are defined using the `Loosened FixedCutLoose` isolation in the muon measurement region with $N_{\text{b-jets}} = 0$.



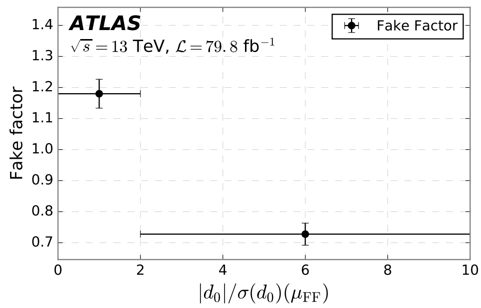
(a) Bin: 10 GeV $\leq p_T(\mu_{\mathrm{FF}}) <$ 20 GeV and $N_{\mathrm{jets}} \leq 2$.



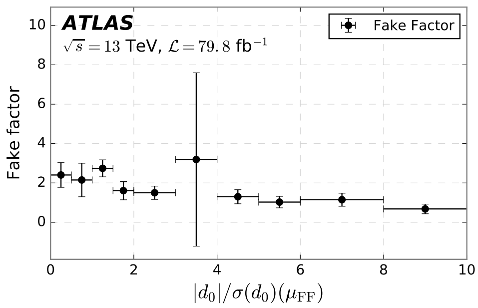
(b) Applied binning as a function of $|d_0|/\sigma(d_0)(\mu_{\mathrm{FF}})$ for Fig. 111a.



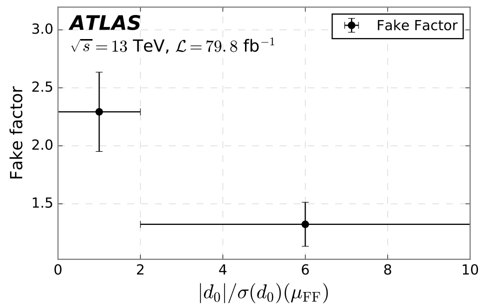
(c) Bin: 20 GeV $\leq p_T(\mu_{\mathrm{FF}}) <$ 30 GeV and $N_{\mathrm{jets}} \leq 2$.



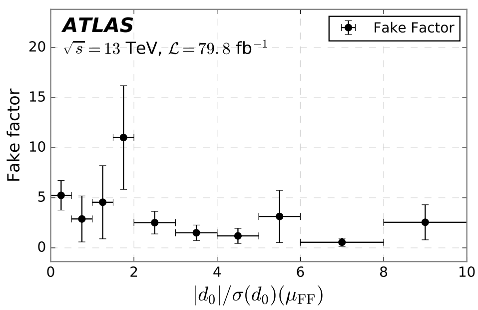
(d) Applied binning as a function of $|d_0|/\sigma(d_0)(\mu_{\mathrm{FF}})$ for Fig. 111c.



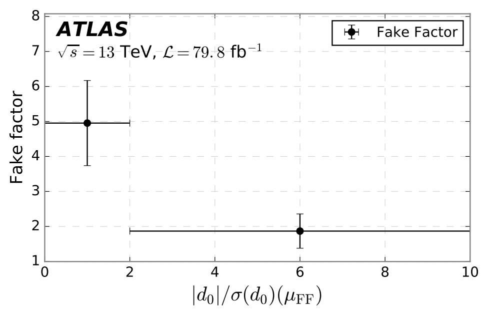
(e) Bin: 30 GeV $\leq p_T(\mu_{\mathrm{FF}}) <$ 40 GeV and $N_{\mathrm{jets}} \leq 2$.



(f) Applied binning as a function of $|d_0|/\sigma(d_0)(\mu_{\mathrm{FF}})$ for Fig. 111e.

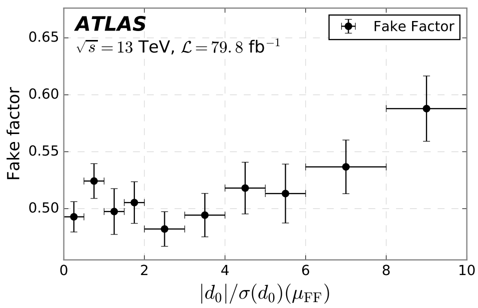


(g) Bin: 40 GeV $\leq p_T(\mu_{\mathrm{FF}}) <$ 50 GeV and $N_{\mathrm{jets}} \leq 2$.

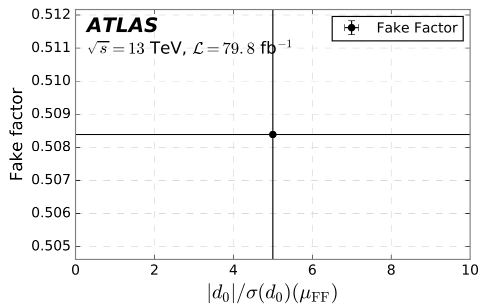


(h) Applied binning as a function of $|d_0|/\sigma(d_0)(\mu_{\mathrm{FF}})$ for Fig. 111g.

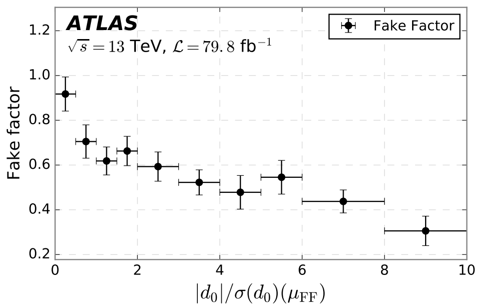
Figure 111: Muon fake factor as a function of $|d_0|/\sigma(d_0)(\mu_{\mathrm{FF}})$. The considered fake factors are binned in $p_T(\mu_{\mathrm{FF}})$ and N_{jets}. The ID and anti-ID criteria are defined using the `Loosened FixedCutLoose` isolation selection in the muon measurement region with $N_{\text{b-jets}} = 0$.



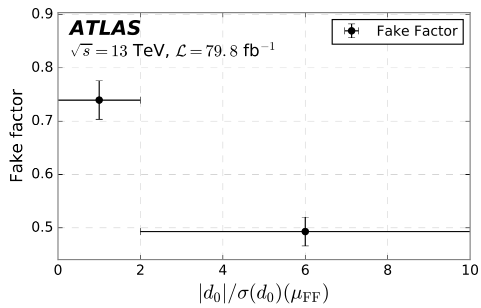
(a) Bin: 10 GeV $\leq p_T(\mu_{\mathrm{FF}}) < 20$ GeV and $N_{\mathrm{jets}} > 2$.



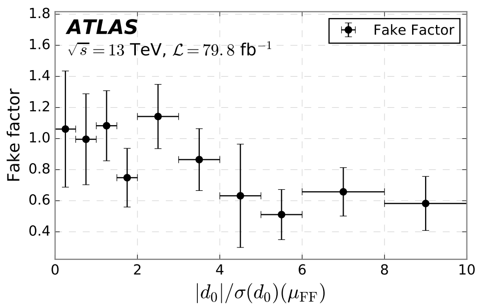
(b) Applied binning as a function of $|d_0|/\sigma(d_0)(\mu_{\mathrm{FF}})$ for Fig. 112a.



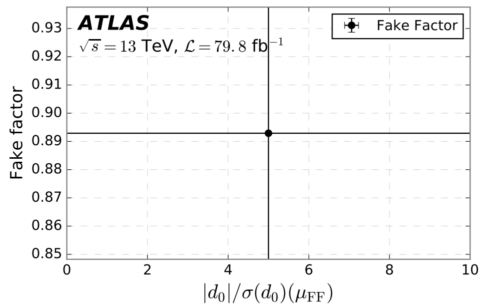
(c) Bin: 20 GeV $\leq p_T(\mu_{\mathrm{FF}}) < 30$ GeV and $N_{\mathrm{jets}} > 2$.



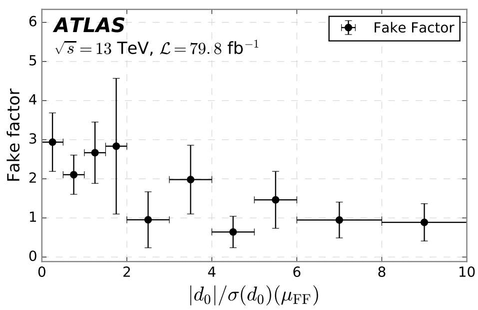
(d) Applied binning as a function of $|d_0|/\sigma(d_0)(\mu_{\mathrm{FF}})$ for Fig. 112c.



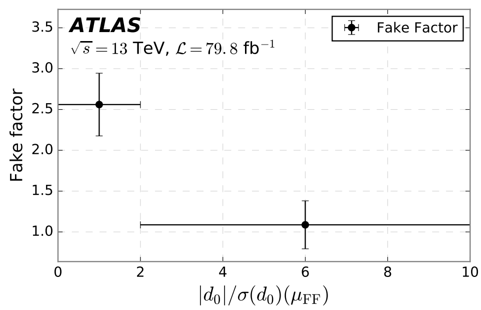
(e) Bin: 30 GeV $\leq p_T(\mu_{\mathrm{FF}}) < 40$ GeV and $N_{\mathrm{jets}} > 2$.



(f) Applied binning as a function of $|d_0|/\sigma(d_0)(\mu_{\mathrm{FF}})$ for Fig. 112e.

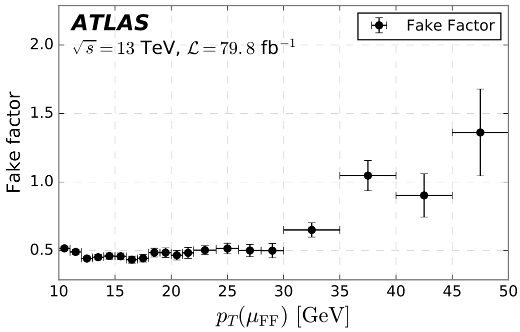


(g) Bin: 40 GeV $\leq p_T(\mu_{\mathrm{FF}}) < 50$ GeV and $N_{\mathrm{jets}} > 2$.

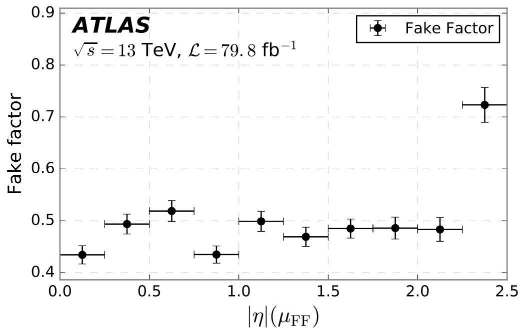


(h) Applied binning as a function of $|d_0|/\sigma(d_0)(\mu_{\mathrm{FF}})$ for Fig. 112g.

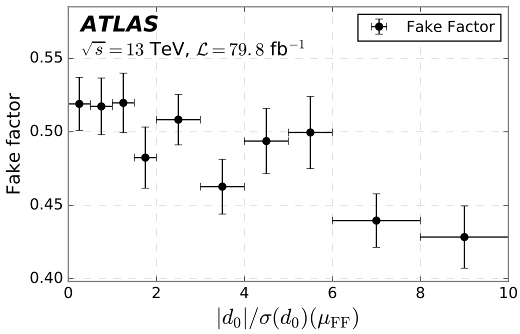
Figure 112: Muon fake factor as a function of $|d_0|/\sigma(d_0)(\mu_{\mathrm{FF}})$. The considered fake factors are binned in $p_T(\mu_{\mathrm{FF}})$ and N_{jets}. The ID and anti-ID criteria are defined using the `Loosened FixedCutLoose` isolation selection in the muon measurement region with $N_{\text{b-jets}} = 0$.



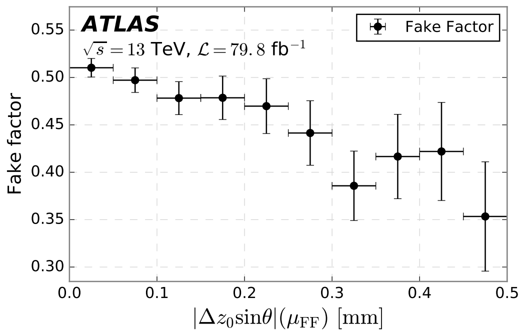
(a) Fake factor as a function of $p_T(\mu_{\mathrm{FF}})$.



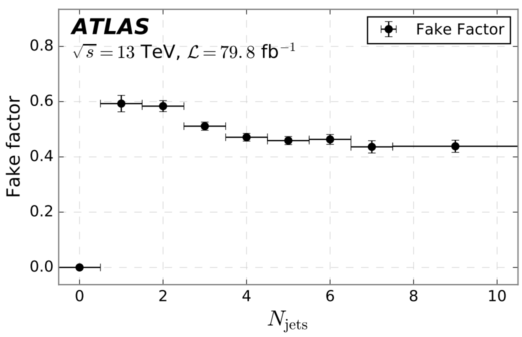
(b) Fake factor as a function of $|\eta|(\mu_{\mathrm{FF}})$.



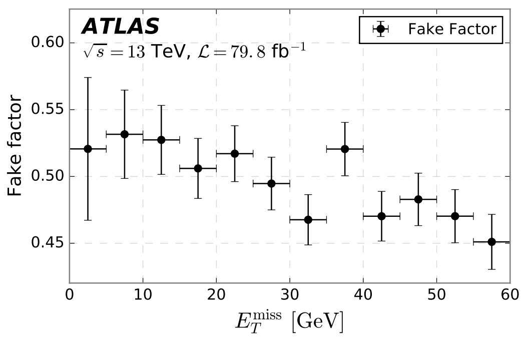
(c) Fake factor as a function of $|d_0|/\sigma(d_0)(\mu_{\mathrm{FF}})$.



(d) Fake factor as a function of $|\Delta z_0 \sin(\theta)|(\mu_{\mathrm{FF}})$.

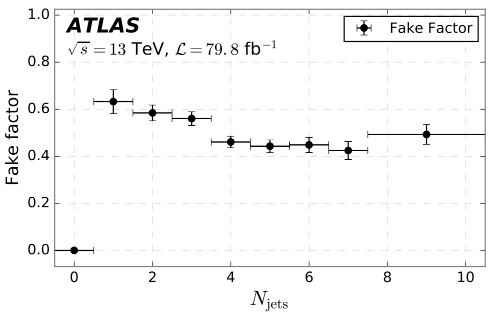


(e) Fake factor as a function of N_{jets}.

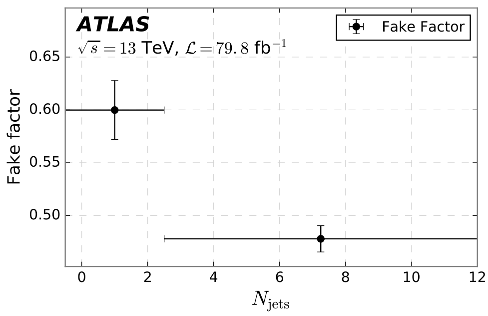


(f) Fake factor as a function of E_T^{miss}.

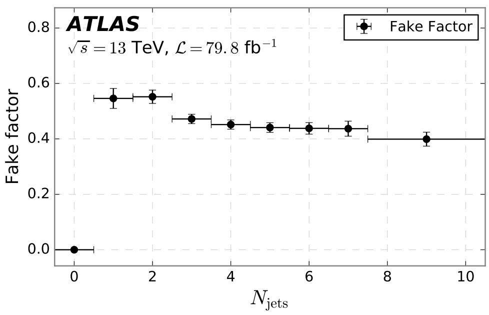
Figure 113: Muon fake factor as a function of: $p_T(\mu_{\mathrm{FF}})$; $|\eta|(\mu_{\mathrm{FF}})$; $|d_0|/\sigma(d_0)(\mu_{\mathrm{FF}})$; $|\Delta z_0 \sin(\theta)|(\mu_{\mathrm{FF}})$; N_{jets}; and E_T^{miss}. The ID and anti-ID criteria are defined using the `Loosened FixedCutLoose` isolation in the muon measurement region with $N_{\text{b-jets}} = 0$.



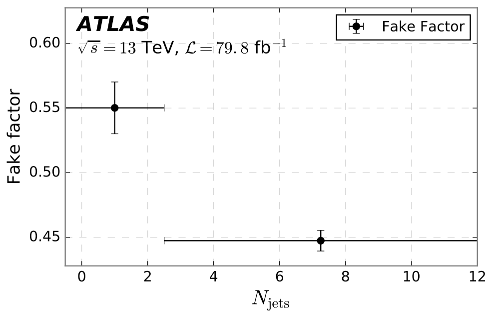
(a) Fake factor as a function of N_{jets}. Bin: 10 GeV $\leq p_T(\mu_{\mathrm{FF}}) <$ 20 GeV.



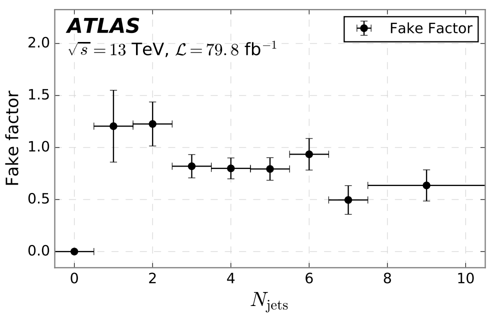
(b) Applied binning as a function of N_{jets} for Fig. 114a.



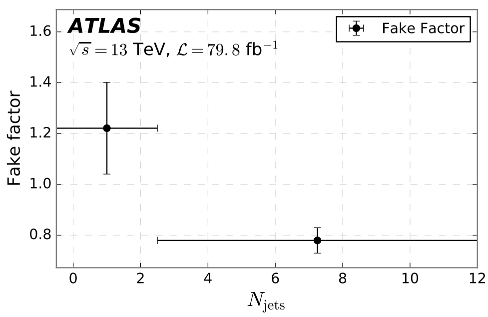
(c) Fake factor as a function of N_{jets}. Bin: 20 GeV $\leq p_T(\mu_{\mathrm{FF}}) <$ 30 GeV.



(d) Applied binning as a function of N_{jets} for Fig. 114c.

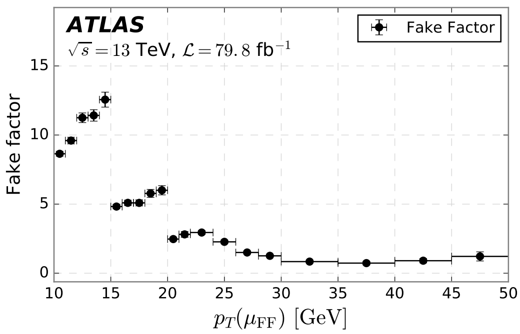


(e) Fake factor as a function of N_{jets}. Bin: 40 GeV $\leq p_T(\mu_{\mathrm{FF}}) <$ 50 GeV.

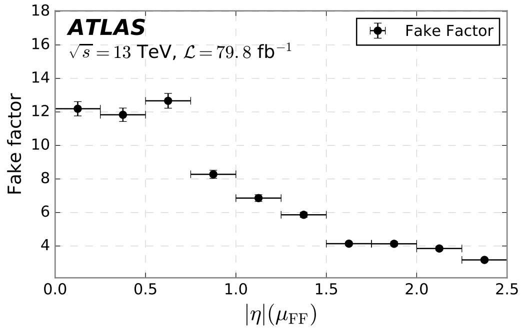


(f) Applied binning as a function of N_{jets} for Fig. 114e.

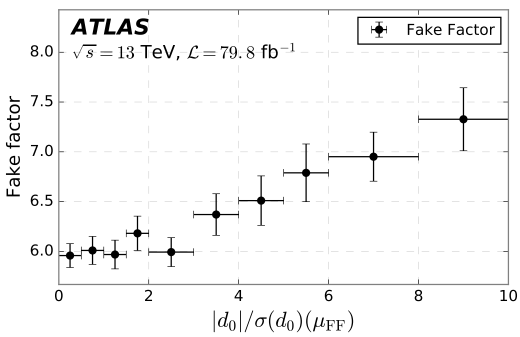
Figure 114: Muon fake factor as a function of N_{jets}. The considered fake factors are binned in $p_T(\mu_{\mathrm{FF}})$. The ID and anti-ID criteria are defined using the `FixedCutLoose` isolation selection in the muon measurement region with $N_{\text{b-jets}} \geq 1$.



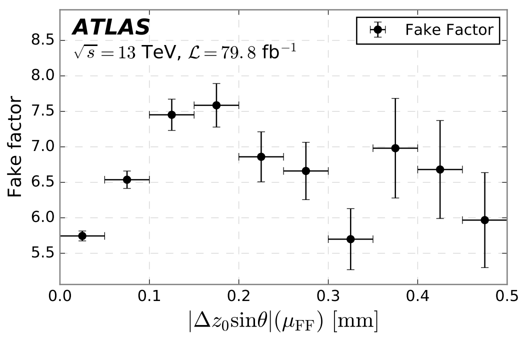
(a) Fake factor as a function of $p_T(\mu_{\text{FF}})$.



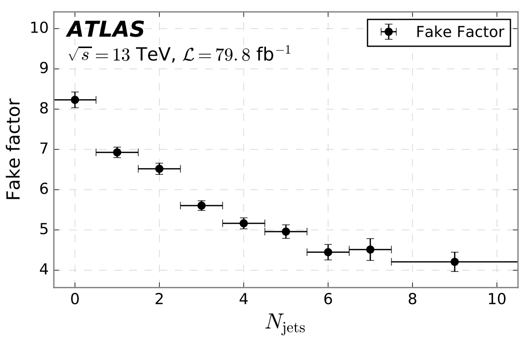
(b) Fake factor as a function of $|\eta|(\mu_{\text{FF}})$.



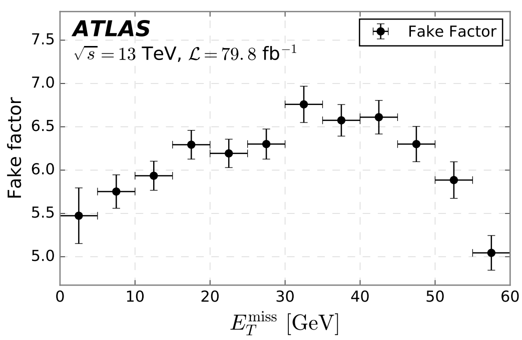
(c) Fake factor as a function of $|d_0|/\sigma(d_0)(\mu_{\text{FF}})$.



(d) Fake factor as a function of $|\Delta z_0 \sin(\theta)|(\mu_{\text{FF}})$.

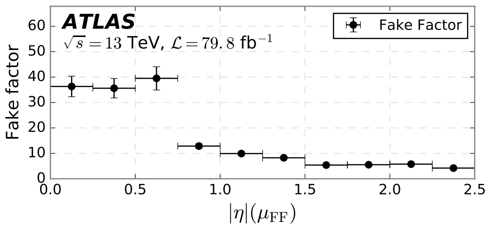


(e) Fake factor as a function of N_{jets}.

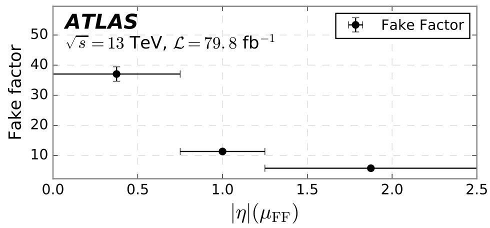


(f) Fake factor as a function of E_T^{miss}.

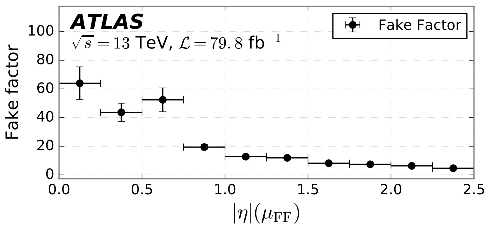
Figure 115: Muon fake factor as a function of: $p_T(\mu_{\text{FF}})$; $|\eta|(\mu_{\text{FF}})$; $|d_0|/\sigma(d_0)(\mu_{\text{FF}})$; $|\Delta z_0 \sin(\theta)|(\mu_{\text{FF}})$; N_{jets}; and E_T^{miss}. The ID and anti-ID criteria are defined using the `Loosene` isolation in the muon measurement region with $N_{\text{b-jets}} = 0$.



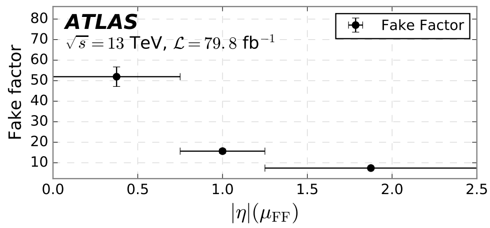
(a) Bin: 10 GeV $\leq p_T(\mu_{\rm FF}) < 20$ GeV and $N_{\rm jets} \leq 2$.



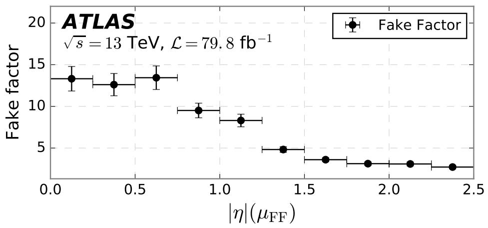
(b) Applied binning as a function of $|\eta|(\mu_{\rm FF})$ for Fig. 116a.



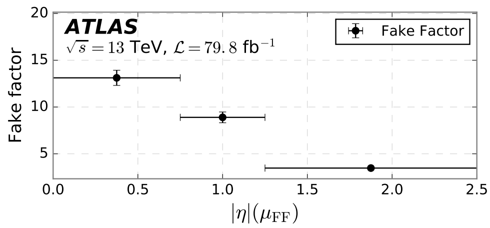
(c) Bin: 10 GeV $\leq p_T(\mu_{\rm FF}) < 20$ GeV and $N_{\rm jets} \leq 2$.



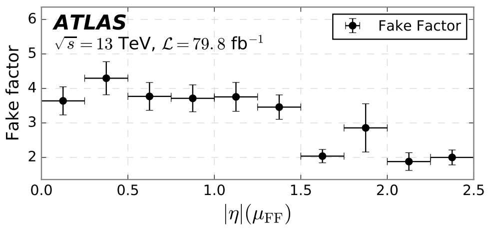
(d) Applied binning as a function of $|\eta|(\mu_{\rm FF})$ for Fig. 116c.



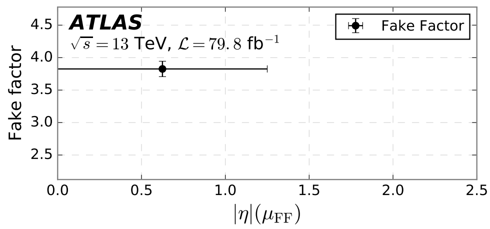
(e) Bin: 10 GeV $\leq p_T(\mu_{\rm FF}) < 20$ GeV and $N_{\rm jets} \leq 2$.



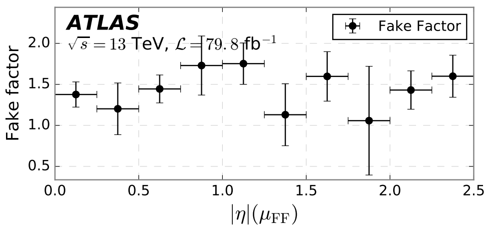
(f) Applied binning as a function of $|\eta|(\mu_{\rm FF})$ for Fig. 116e.



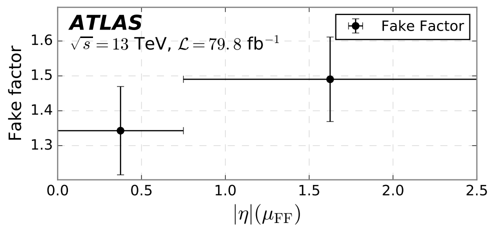
(g) Bin: 10 GeV $\leq p_T(\mu_{\rm FF}) < 20$ GeV and $N_{\rm jets} \leq 2$.



(h) Applied binning as a function of $|\eta|(\mu_{\rm FF})$ for Fig. 116g.

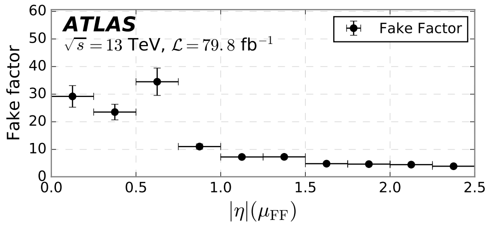


(i) Bin: 10 GeV $\leq p_T(\mu_{\rm FF}) < 20$ GeV and $N_{\rm jets} \leq 2$.

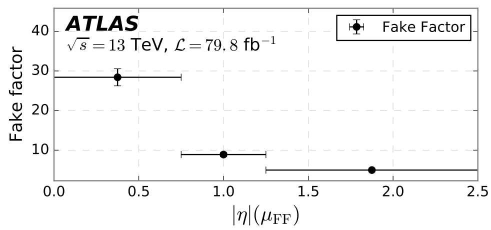


(j) Applied binning as a function of $|\eta|(\mu_{\rm FF})$ for Fig. 116i.

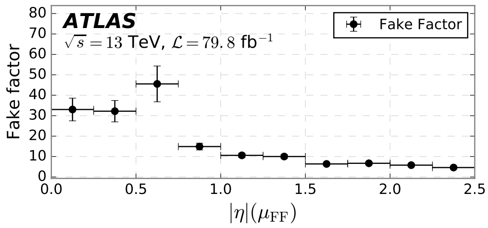
Figure 116: Muon fake factor as a function of $|\eta|(\mu_{\rm FF})$. The considered fake factors are binned in $p_T(\mu_{\rm FF})$ and $N_{\rm jets}$. The ID and anti-ID criteria are defined using the `Loose` isolation selection in the muon measurement region with $N_{\text{b-jets}} = 0$.



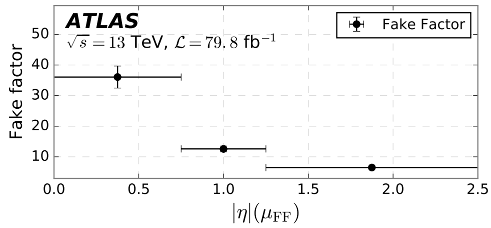
(a) Bin: 10 GeV $\leq p_T(\mu_{\mathrm{FF}}) <$ 20 GeV and $N_{\mathrm{jets}} > 2$.



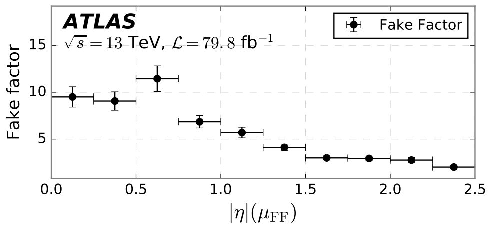
(b) Applied binning as a function of $|\eta|(\mu_{\mathrm{FF}})$ for Fig. 117a.



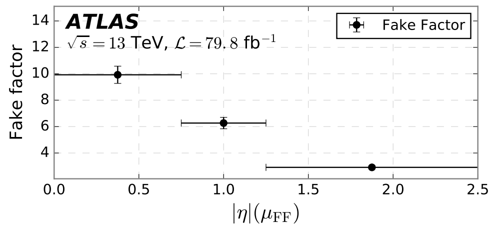
(c) Bin: 10 GeV $\leq p_T(\mu_{\mathrm{FF}}) <$ 20 GeV and $N_{\mathrm{jets}} > 2$.



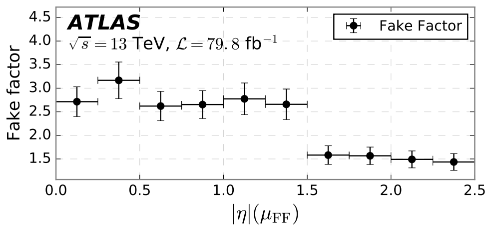
(d) Applied binning as a function of $|\eta|(\mu_{\mathrm{FF}})$ for Fig. 117c.



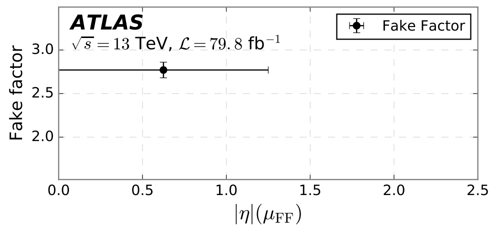
(e) Bin: 10 GeV $\leq p_T(\mu_{\mathrm{FF}}) <$ 20 GeV and $N_{\mathrm{jets}} > 2$.



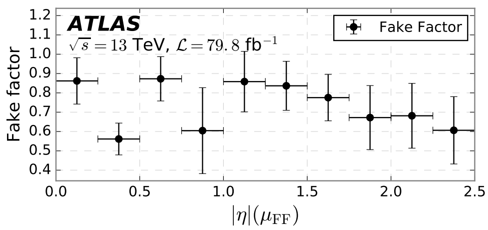
(f) Applied binning as a function of $|\eta|(\mu_{\mathrm{FF}})$ for Fig. 117e.



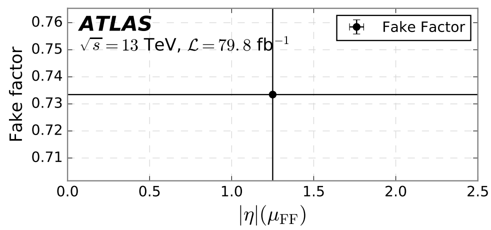
(g) Bin: 10 GeV $\leq p_T(\mu_{\mathrm{FF}}) <$ 20 GeV and $N_{\mathrm{jets}} > 2$.



(h) Applied binning as a function of $|\eta|(\mu_{\mathrm{FF}})$ for Fig. 117g.

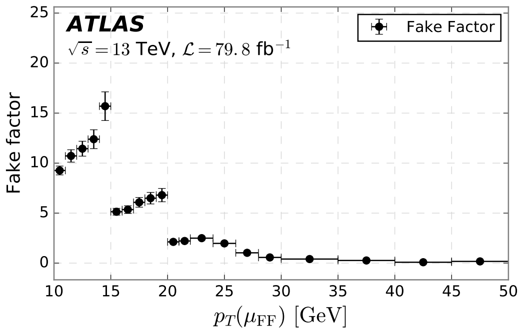


(i) Bin: 10 GeV $\leq p_T(\mu_{\mathrm{FF}}) <$ 20 GeV and $N_{\mathrm{jets}} > 2$.

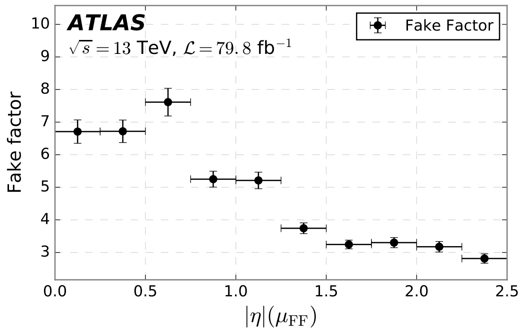


(j) Applied binning as a function of $|\eta|(\mu_{\mathrm{FF}})$ for Fig. 117i.

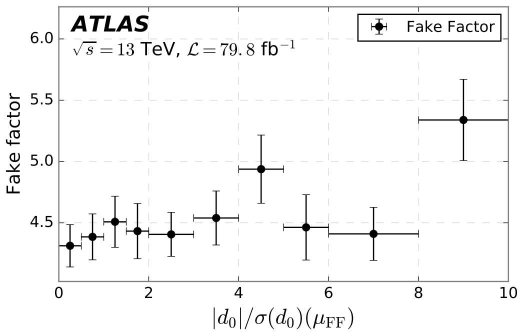
Figure 117: Muon fake factor as a function of $|\eta|(\mu_{\mathrm{FF}})$. The considered fake factors are binned in $p_T(\mu_{\mathrm{FF}})$ and N_{jets}. The ID and anti-ID criteria are defined using the `Loose` isolation selection in the muon measurement region with $N_{\text{b-jets}} = 0$.



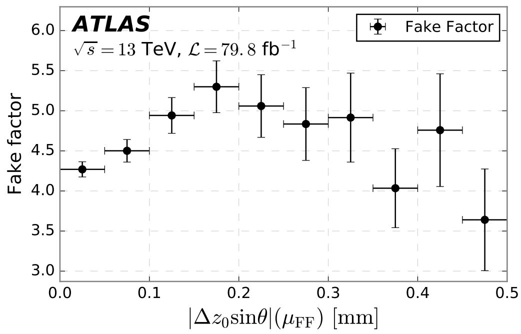
(a) Fake factor as a function of $p_T(\mu_{\mathrm{FF}})$.



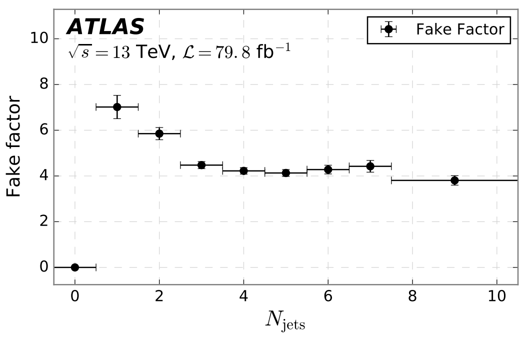
(b) Fake factor as a function of $|\eta|(\mu_{\mathrm{FF}})$.



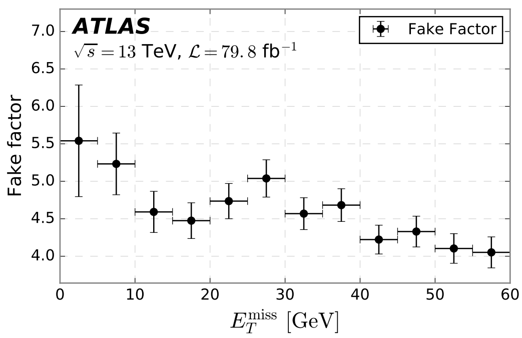
(c) Fake factor as a function of $|d_0|/\sigma(d_0)(\mu_{\mathrm{FF}})$.



(d) Fake factor as a function of $|\Delta z_0 \sin(\theta)|(\mu_{\mathrm{FF}})$.

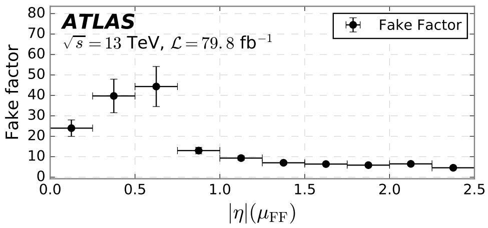


(e) Fake factor as a function of N_{jets}.

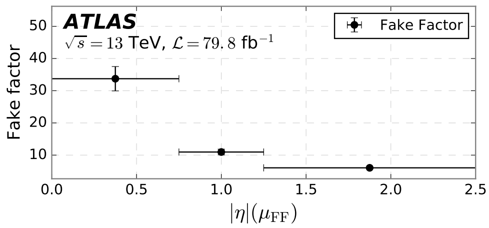


(f) Fake factor as a function of E_T^{miss}.

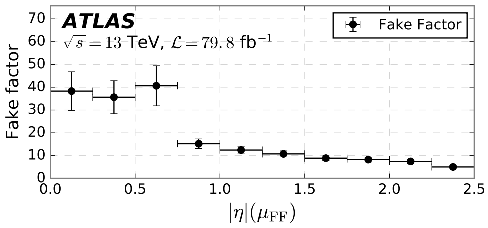
Figure 118: Muon fake factor as a function of: $p_T(\mu_{\mathrm{FF}})$; $|\eta|(\mu_{\mathrm{FF}})$; $|d_0|/\sigma(d_0)(\mu_{\mathrm{FF}})$; $|\Delta z_0 \sin(\theta)|(\mu_{\mathrm{FF}})$; N_{jets}; and E_T^{miss}. The ID and anti-ID criteria are defined using the `Loosene` isolation in the muon measurement region with $N_{\text{b-jets}} = 0$.



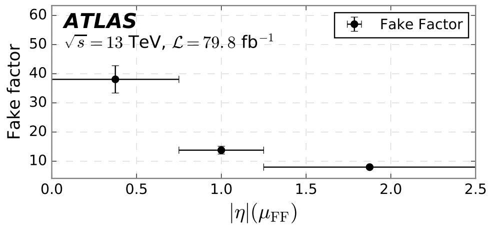
(a) Fake factor as a function of N_{jets}. Bin: 10 GeV $\leq p_T(\mu_{\text{FF}}) <$ 12 GeV.



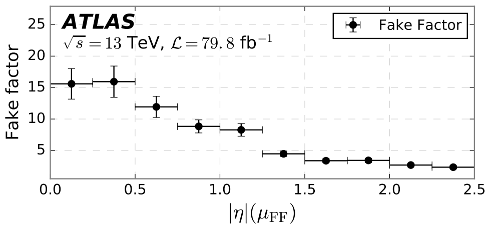
(b) Applied binning as a function of N_{jets} for Fig. 119a.



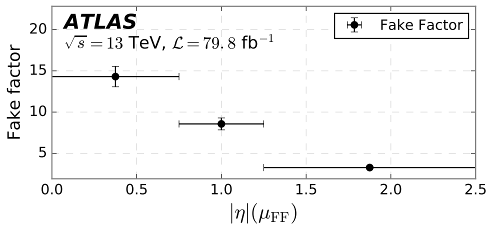
(c) Fake factor as a function of N_{jets}. Bin: 12 GeV $\leq p_T(\mu_{\text{FF}}) <$ 15 GeV.



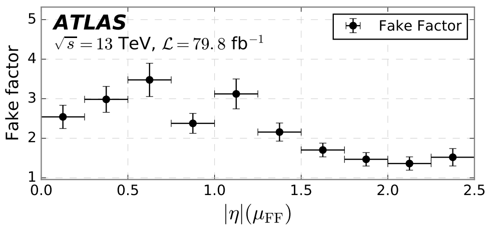
(d) Applied binning as a function of N_{jets} for Fig. 119c.



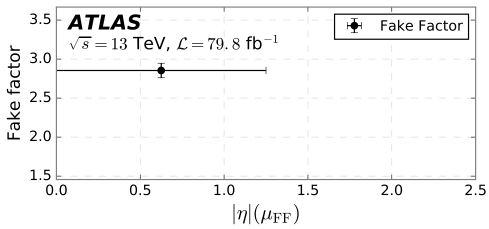
(e) Fake factor as a function of N_{jets}. Bin: 15 GeV $\leq p_T(\mu_{\text{FF}}) <$ 20 GeV.



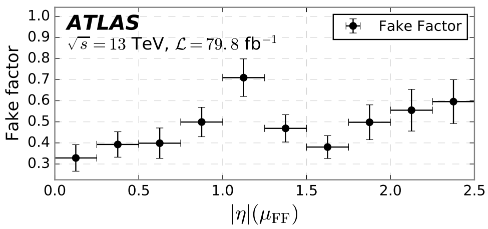
(f) Applied binning as a function of N_{jets} for Fig. 119e.



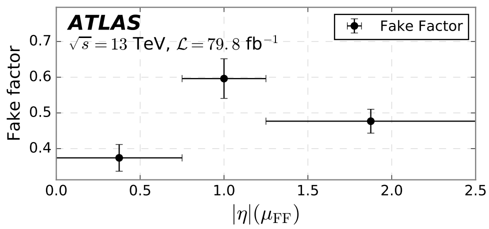
(g) Fake factor as a function of N_{jets}. Bin: 20 GeV $\leq p_T(\mu_{\text{FF}}) <$ 26 GeV.



(h) Applied binning as a function of N_{jets} for Fig. 119g.

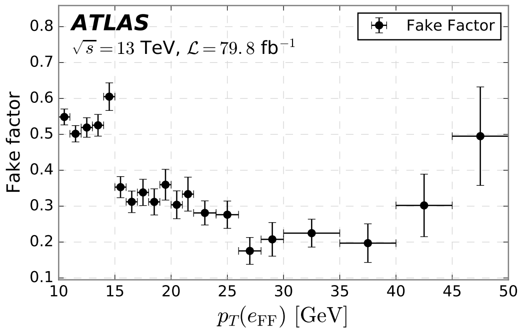


(i) Fake factor as a function of N_{jets}. Bin: 26 GeV $\leq p_T(\mu_{\text{FF}}) <$ 50 GeV.

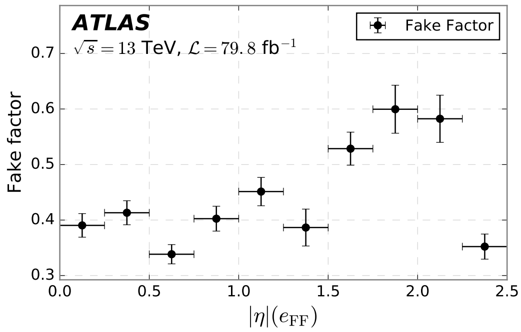


(j) Applied binning as a function of N_{jets} for Fig. 119i.

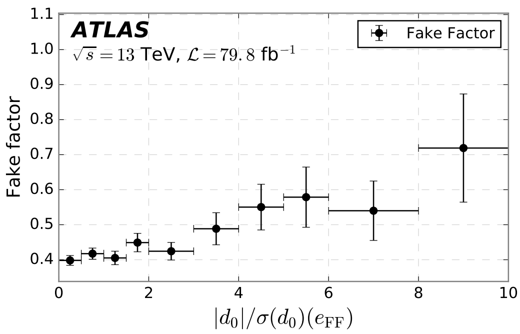
Figure 119: Muon fake factor as a function of N_{jets}. The considered fake factors are binned in $p_T(\mu_{\text{FF}})$. The ID and anti-ID criteria are defined using the `FixedCutLoose` isolation selection in the muon measurement region with $N_{\text{b-jets}} \geq 1$.



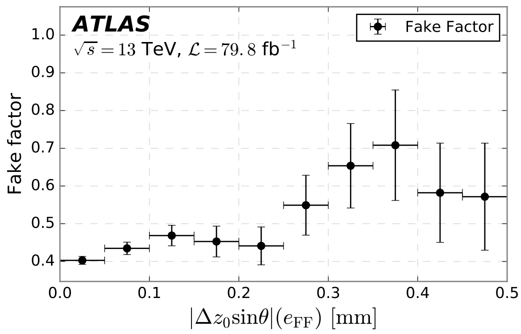
(a) Fake factor as a function of $p_T(e_{\mathrm{FF}})$.



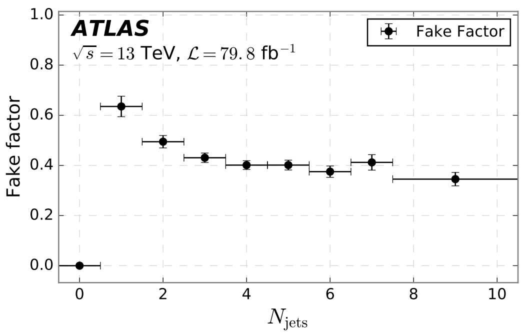
(b) Fake factor as a function of $|\eta|(e_{\mathrm{FF}})$.



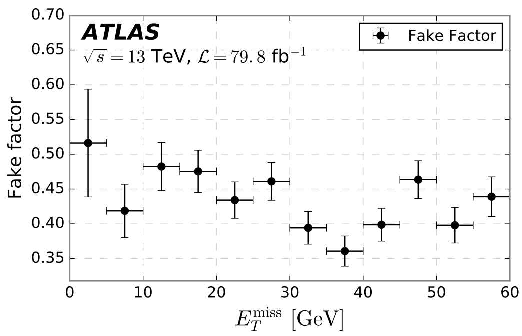
(c) Fake factor as a function of $|d_0|/\sigma(d_0)(e_{\mathrm{FF}})$.



(d) Fake factor as a function of $|\Delta z_0 \sin(\theta)|(e_{\mathrm{FF}})$.

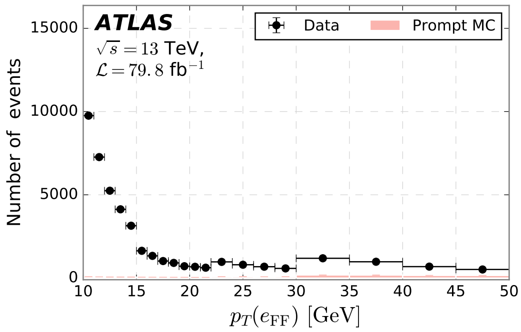


(e) Fake factor as a function of N_{jets}.

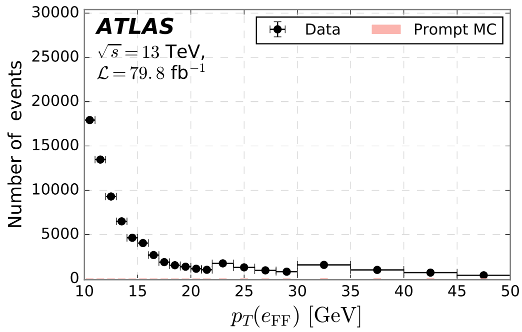


(f) Fake factor as a function of E_T^{miss}.

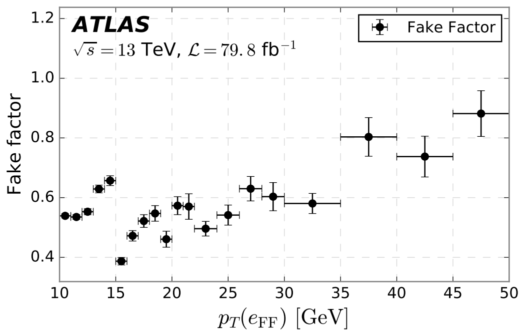
Figure 120: Electron fake factor as a function of: $p_T(e_{\mathrm{FF}})$; $|\eta|(e_{\mathrm{FF}})$; $|d_0|/\sigma(d_0)(e_{\mathrm{FF}})$; $|\Delta z_0 \sin(\theta)|(e_{\mathrm{FF}})$; N_{jets}; and E_T^{miss}. The ID and anti-ID criteria are defined using the `Gradient` isolation in the electron measurement region with $N_{\text{b-jets}} \geq 1$.



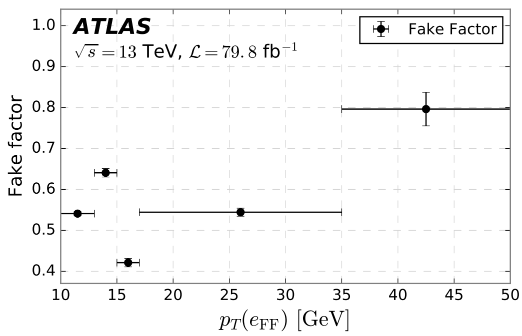
(a) Fake factor numerator: ID for e_{FF}. Prompt MC: Sum of all MC background processes. e_{FF} is required to be real.



(b) Fake factor denominator: anti-ID for e_{FF}. Prompt MC: Sum of all MC background processes. e_{FF} is required to be real.

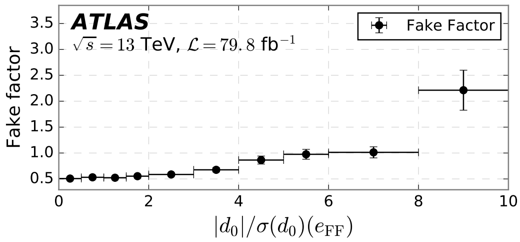


(c) Fake factor as a function of $p_T(e_{\mathrm{FF}})$.

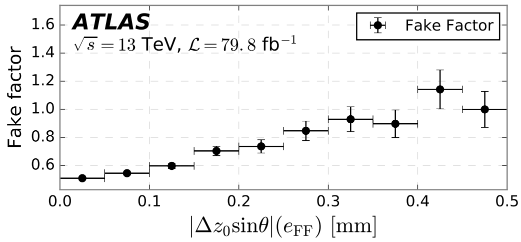


(d) Applied binning as a function of $p_T(e_{\mathrm{FF}})$ for Fig. 121c.

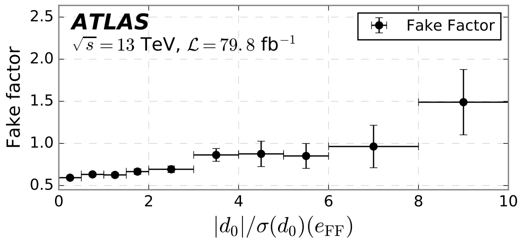
Figure 121: Electron fake factor as a function of $p_T(e_{\mathrm{FF}})$. The ID and anti-ID criteria are defined using the `Gradient` isolation in the electron measurement region with $N_{\text{b-jets}} = 0$.



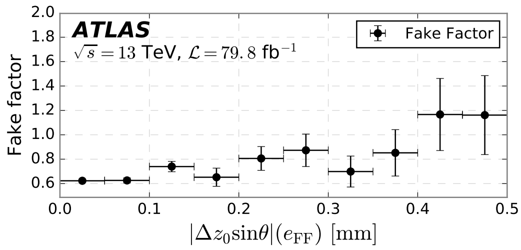
(a) Distribution: $|d_0|/\sigma(d_0)$. Bin 1 in p_T.



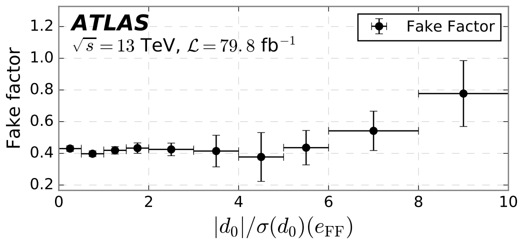
(b) Distribution: $|\Delta z_0 \sin(\theta)|$. Bin 1 in p_T.



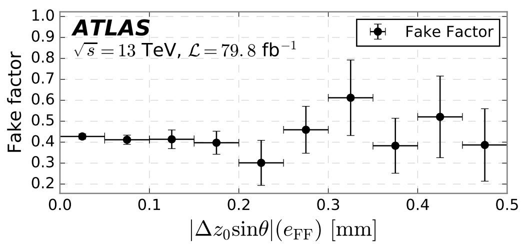
(c) Distribution: $|d_0|/\sigma(d_0)$. Bin 2 in p_T.



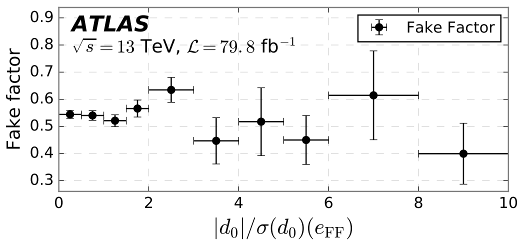
(d) Distribution: $|\Delta z_0 \sin(\theta)|$. Bin 2 in p_T.



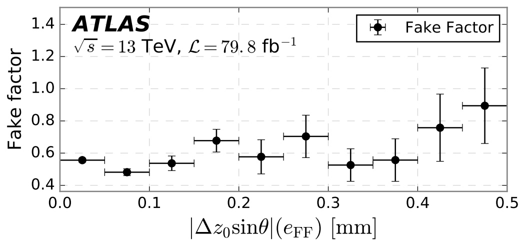
(e) Distribution: $|d_0|/\sigma(d_0)$. Bin 3 in p_T.



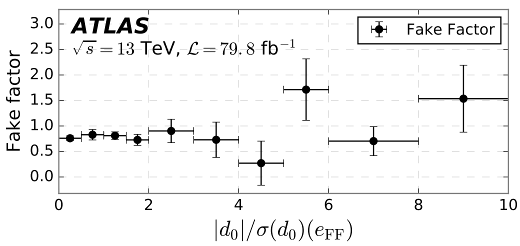
(f) Distribution: $|\Delta z_0 \sin(\theta)|$. Bin 3 in p_T.



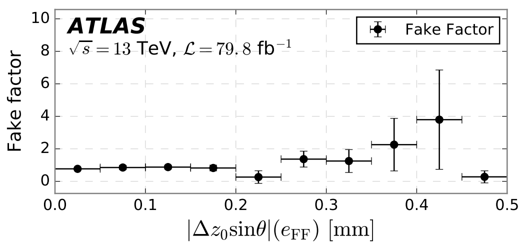
(g) Distribution: $|d_0|/\sigma(d_0)$. Bin 4 in p_T.



(h) Distribution: $|\Delta z_0 \sin(\theta)|$. Bin 4 in p_T.

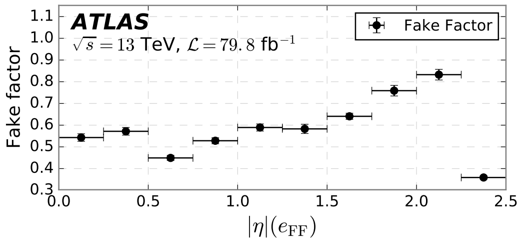


(i) Distribution: $|d_0|/\sigma(d_0)$. Bin 5 in p_T.

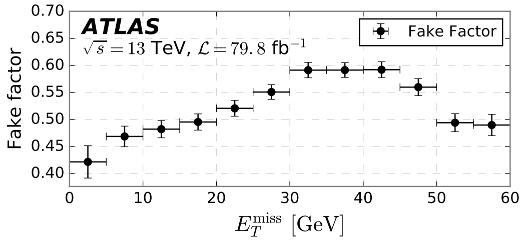


(j) Distribution: $|\Delta z_0 \sin(\theta)|$. Bin 5 in p_T.

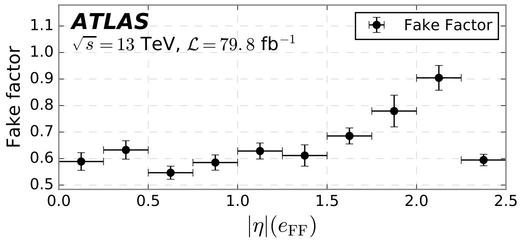
Figure 122: Electron fake factor as a function of $|d_0|/\sigma(d_0)(e_{\mathrm{FF}})$ and $|\Delta z_0 \sin(\theta)|(e_{\mathrm{FF}})$. The considered fake factors are binned in $p_T(\mu_{\mathrm{FF}})$. The ID and anti-ID criteria are defined using the `Gradient` isolation in the electron measurement region with $N_{\text{b-jets}} = 0$.



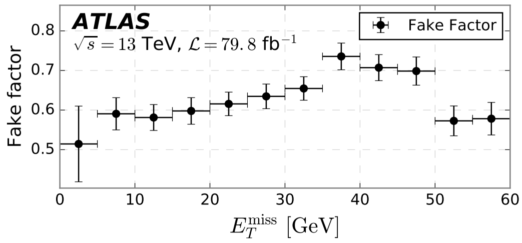
(a) Distribution: $|\eta|$. Bin 1 in p_T.



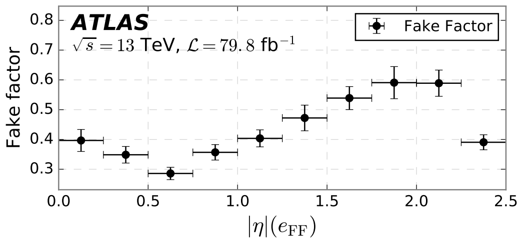
(b) Distribution: E_T^{miss}. Bin 1 in p_T.



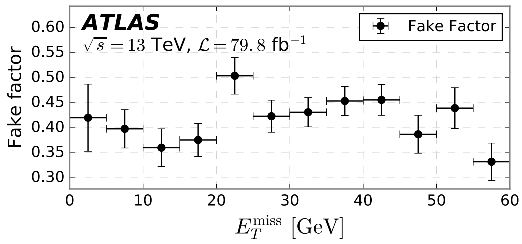
(c) Distribution: $|\eta|$. Bin 2 in p_T.



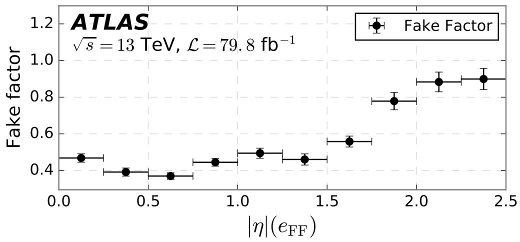
(d) Distribution: E_T^{miss}. Bin 2 in p_T.



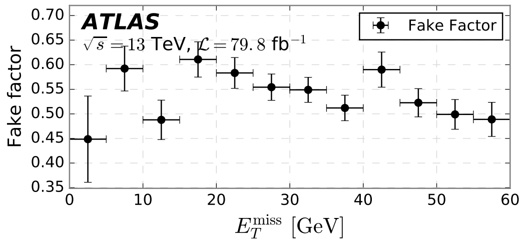
(e) Distribution: $|\eta|$. Bin 3 in p_T.



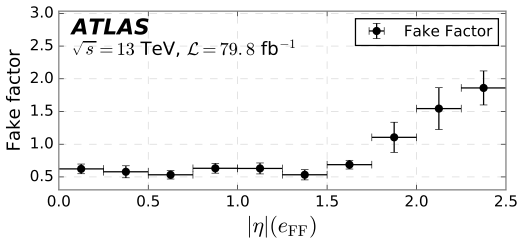
(f) Distribution: E_T^{miss}. Bin 3 in p_T.



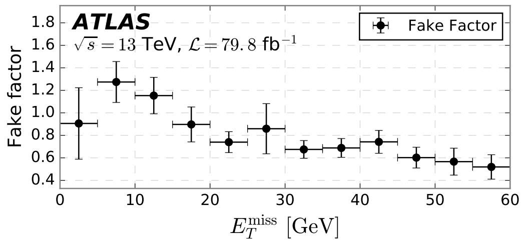
(g) Distribution: $|\eta|$. Bin 4 in p_T.



(h) Distribution: E_T^{miss}. Bin 4 in p_T.

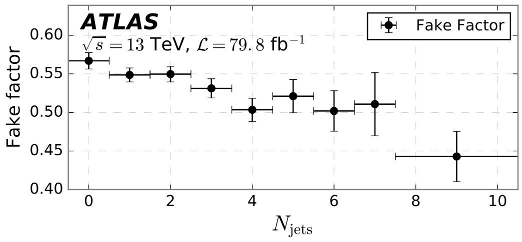


(i) Distribution: $|\eta|$. Bin 5 in p_T.

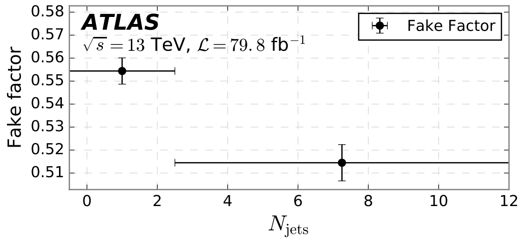


(j) Distribution: E_T^{miss}. Bin 5 in p_T.

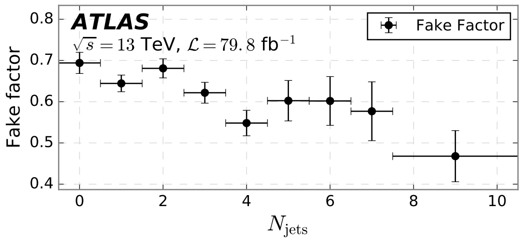
Figure 123: Electron fake factor as a function of $|\eta|(e_{\mathrm{FF}})$ and E_T^{miss}. The considered fake factors are binned in $p_T(\mu_{\mathrm{FF}})$. The ID and anti-ID criteria are defined using the `Gradient` isolation in the electron measurement region with $N_{\text{b-jets}} = 0$.



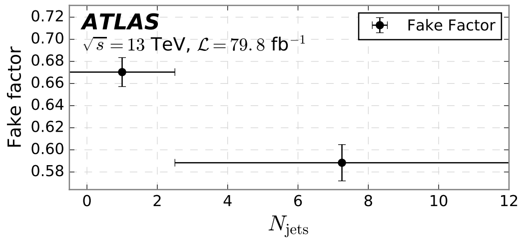
(a) Distribution: N_{jets}. Bin 1 in p_T.



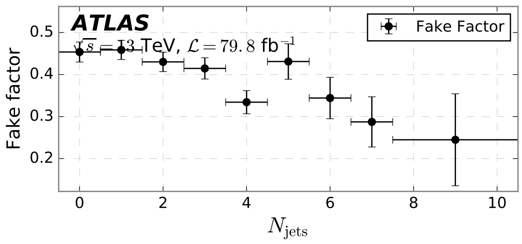
(b) Applied binning for Fig. 124a.



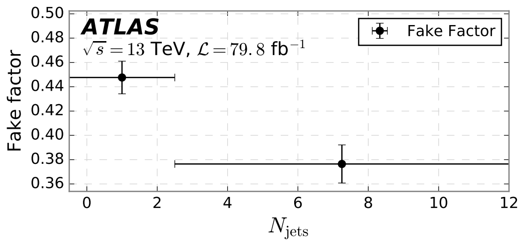
(c) Distribution: N_{jets}. Bin 2 in p_T.



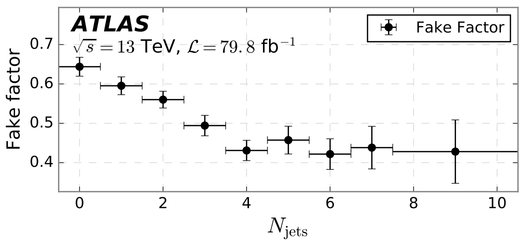
(d) Applied binning for Fig. 124c.



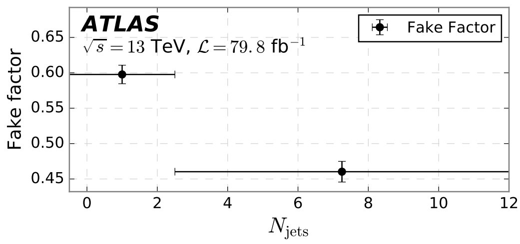
(e) Distribution: N_{jets}. Bin 3 in p_T.



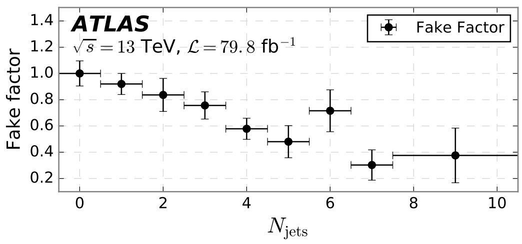
(f) Applied binning for Fig. 124e.



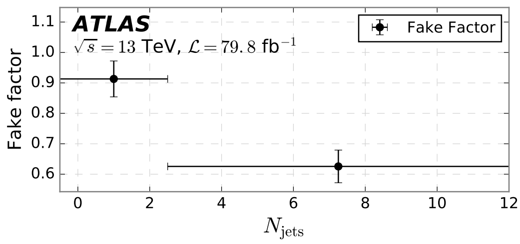
(g) Distribution: N_{jets}. Bin 4 in p_T.



(h) Applied binning for Fig. 124g.

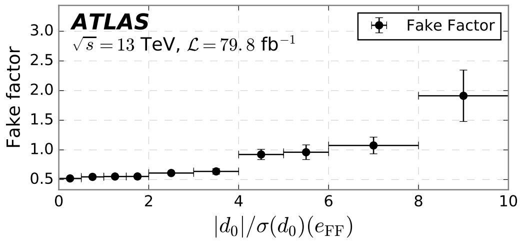


(i) Distribution: N_{jets}. Bin 5 in p_T.

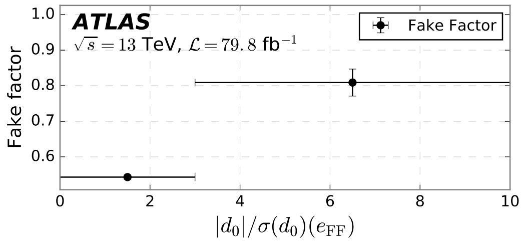


(j) Applied binning for Fig. 124i.

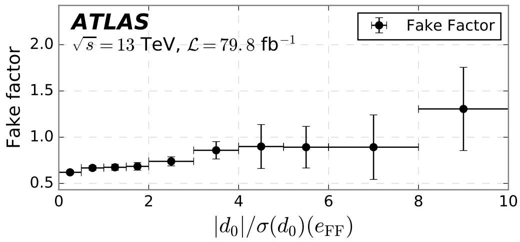
Figure 124: Electron fake factor as a function of N_{jets}. The considered fake factors are binned in $p_T(\mu_{\text{FF}})$. The ID and anti-ID criteria are defined using the `Gradient` isolation in the electron measurement region with $N_{\text{b-jets}} = 0$.



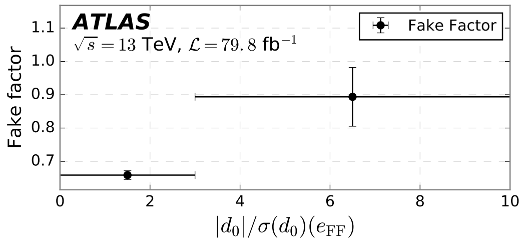
(a) Bin: 10 GeV $\leq p_T(e_{\mathrm{FF}}) <$ 13 GeV and $N_{\mathrm{jets}} \leq 2$.



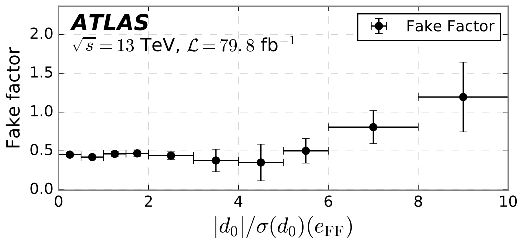
(b) Applied binning for Fig. 125a.



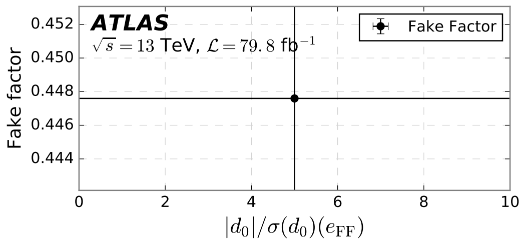
(c) Bin: 13 GeV $\leq p_T(e_{\mathrm{FF}}) <$ 15 GeV and $N_{\mathrm{jets}} \leq 2$.



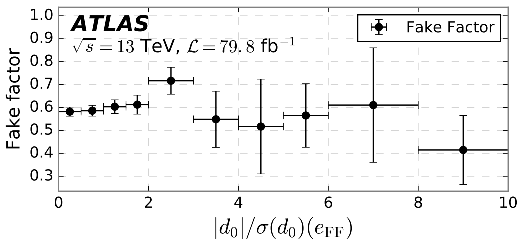
(d) Applied binning for Fig. 125c.



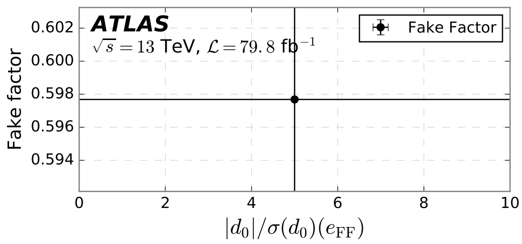
(e) Bin: 15 GeV $\leq p_T(e_{\mathrm{FF}}) <$ 17 GeV and $N_{\mathrm{jets}} \leq 2$.



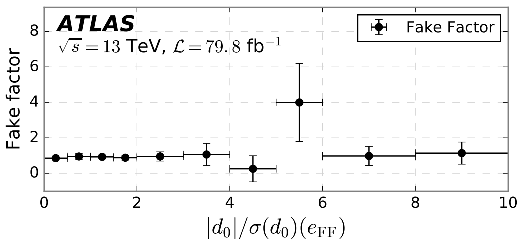
(f) Applied binning for Fig. 125e.



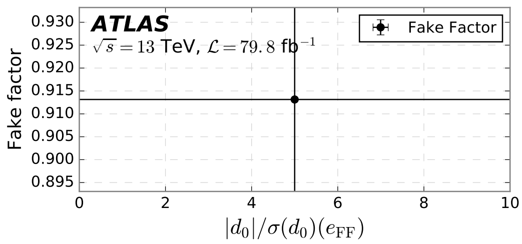
(g) Bin: 17 GeV $\leq p_T(e_{\mathrm{FF}}) <$ 35 GeV and $N_{\mathrm{jets}} \leq 2$.



(h) Applied binning for Fig. 125g.

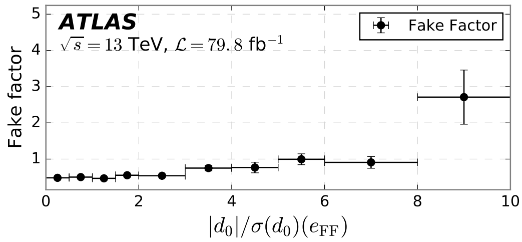


(i) Bin: 35 GeV $\leq p_T(e_{\mathrm{FF}}) <$ 50 GeV and $N_{\mathrm{jets}} \leq 2$.

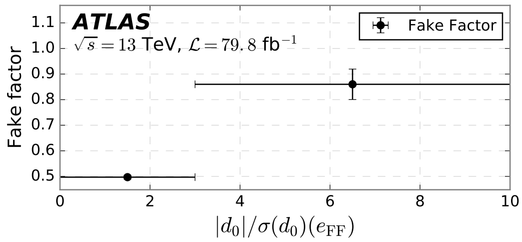


(j) Applied binning for Fig. 125i.

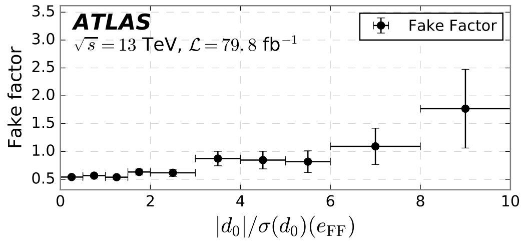
Figure 125: Electron fake factor as a function of $|d_0|/\sigma(d_0)(e_{\mathrm{FF}})$. The considered fake factors are binned in $p_T(\mu_{\mathrm{FF}})$ and N_{jets}. The ID and anti-ID criteria are defined using the `Gradient` isolation in the electron measurement region with $N_{\text{b-jets}} = 0$.



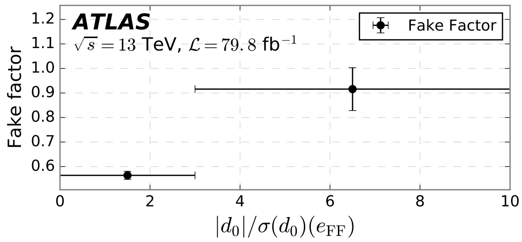
(a) Bin: 10 GeV $\leq p_T(e_{\mathrm{FF}}) < 13$ GeV and $N_{\mathrm{jets}} > 2$.



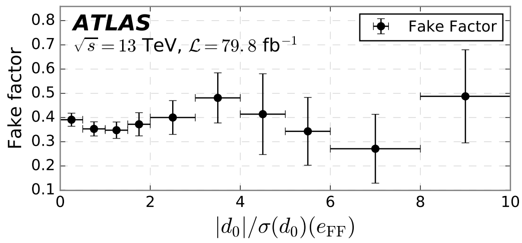
(b) Applied binning for Fig. 126a.



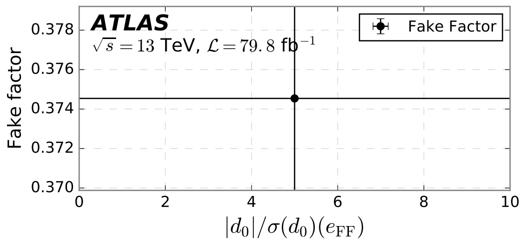
(c) Bin: 13 GeV $\leq p_T(e_{\mathrm{FF}}) < 15$ GeV and $N_{\mathrm{jets}} > 2$.



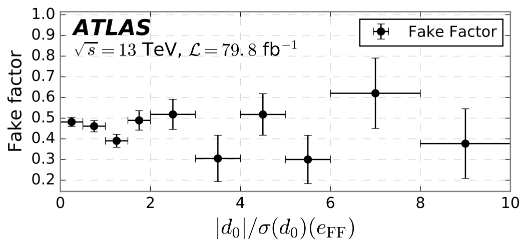
(d) Applied binning for Fig. 126c.



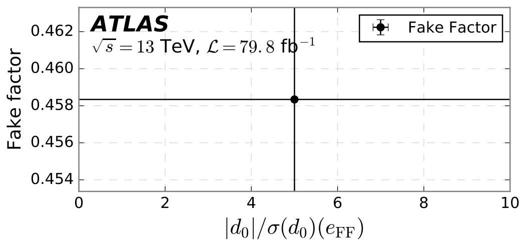
(e) Bin: 15 GeV $\leq p_T(e_{\mathrm{FF}}) < 17$ GeV and $N_{\mathrm{jets}} > 2$.



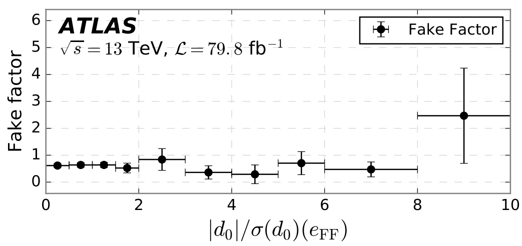
(f) Applied binning for Fig. 126e.



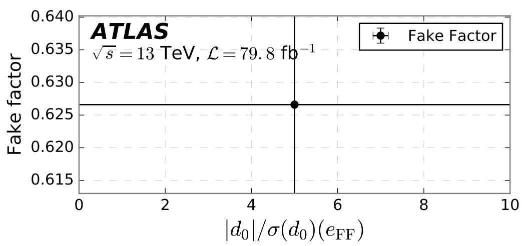
(g) Bin: 17 GeV $\leq p_T(e_{\mathrm{FF}}) < 35$ GeV and $N_{\mathrm{jets}} > 2$.



(h) Applied binning for Fig. 126g.

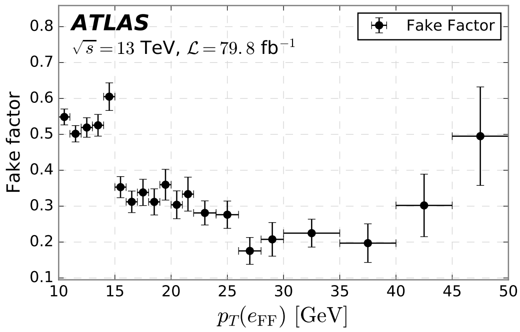


(i) Bin: 35 GeV $\leq p_T(e_{\mathrm{FF}}) < 50$ GeV and $N_{\mathrm{jets}} > 2$.

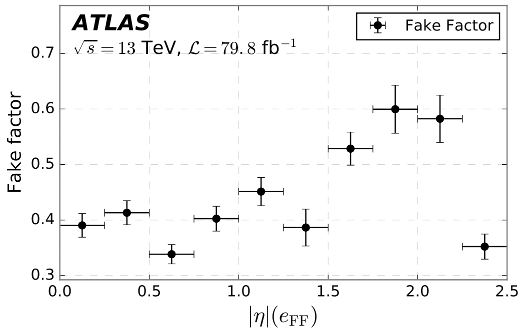


(j) Applied binning for Fig. 126i.

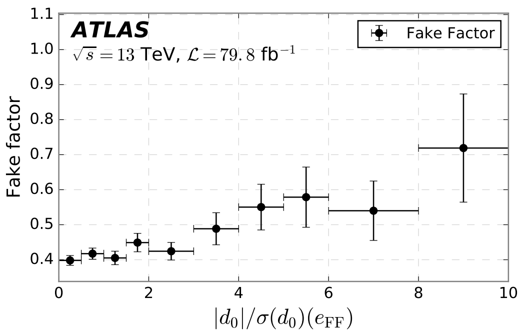
Figure 126: Electron fake factor as a function of $|d_0|/\sigma(d_0)(e_{\mathrm{FF}})$. The considered fake factors are binned in $p_T(\mu_{\mathrm{FF}})$ and N_{jets}. The ID and anti-ID criteria are defined using the `Gradient` isolation in the electron measurement region with $N_{\text{b-jets}} = 0$.



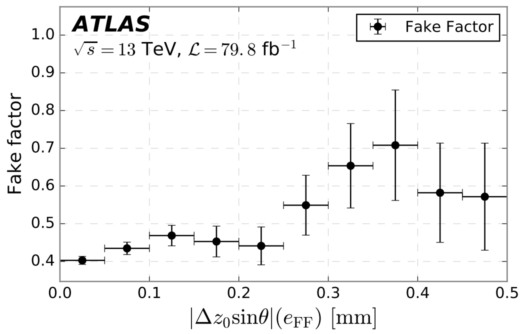
(a) Fake factor as a function of $p_T(e_{\mathrm{FF}})$.



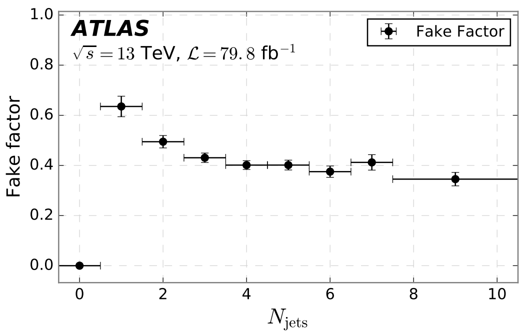
(b) Fake factor as a function of $|\eta|(e_{\mathrm{FF}})$.



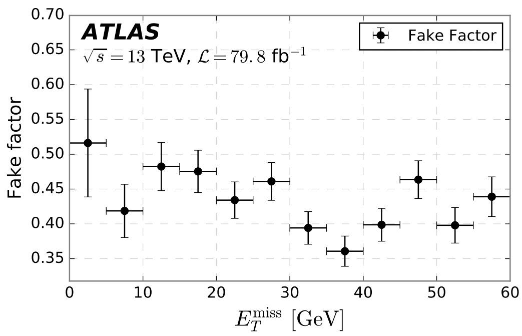
(c) Fake factor as a function of $|d_0|/\sigma(d_0)(e_{\mathrm{FF}})$.



(d) Fake factor as a function of $|\Delta z_0 \sin(\theta)|(e_{\mathrm{FF}})$.

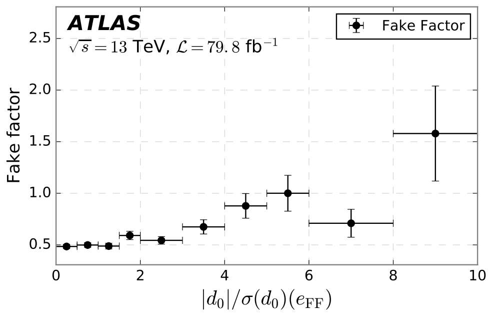


(e) Fake factor as a function of N_{jets}.

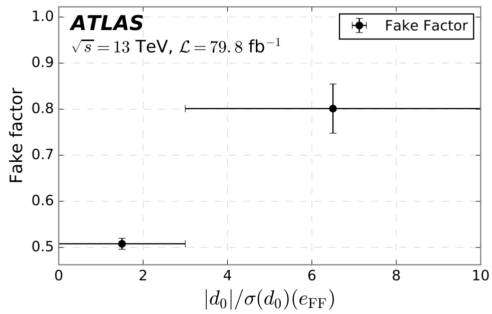


(f) Fake factor as a function of E_T^{miss}.

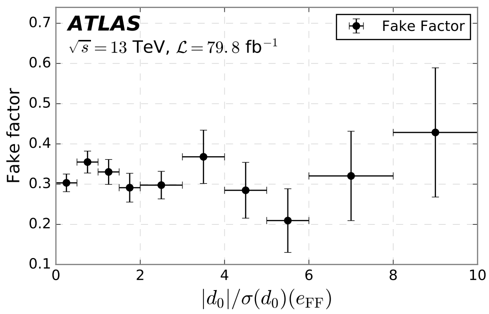
Figure 127: Electron fake factor as a function of: $p_T(e_{\mathrm{FF}})$; $|\eta|(e_{\mathrm{FF}})$; $|d_0|/\sigma(d_0)(e_{\mathrm{FF}})$; $|\Delta z_0 \sin(\theta)|(e_{\mathrm{FF}})$; N_{jets}; and E_T^{miss}. The ID and anti-ID criteria are defined using the `Gradient` isolation in the electron measurement region with $N_{\text{b-jets}} \geq 1$.



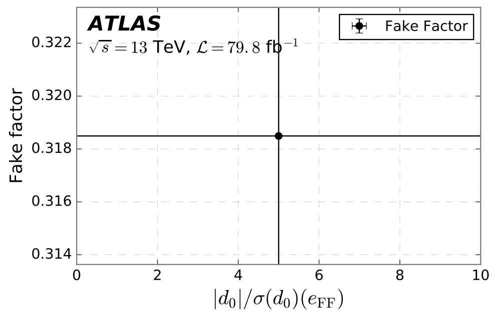
(a) Bin: 10 GeV $\leq p_T(e_{\mathrm{FF}}) <$ 15 GeV.



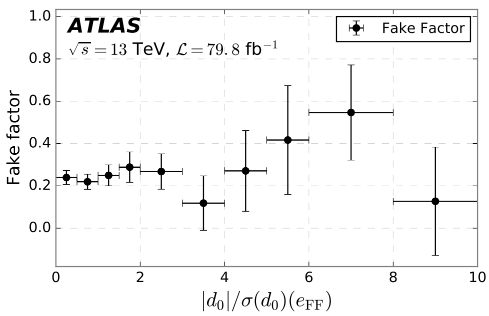
(b) Applied binning for Fig. 128a.



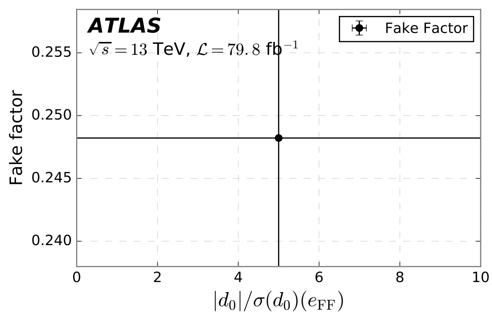
(c) Bin: 15 GeV $\leq p_T(e_{\mathrm{FF}}) <$ 25 GeV.



(d) Applied binning for Fig. 128c.

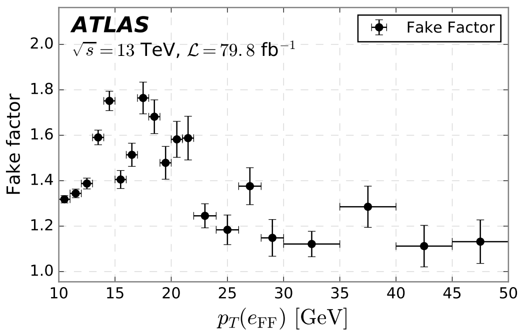


(e) Bin: 25 GeV $\leq p_T(e_{\mathrm{FF}}) <$ 50 GeV.

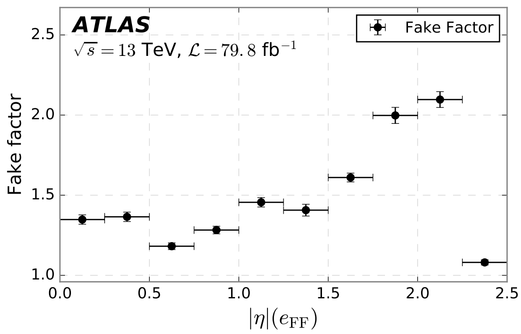


(f) Applied binning for Fig. 128e.

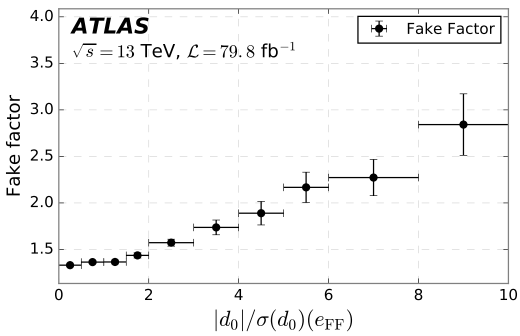
Figure 128: Electron fake factor as a function of $|d_0|/\sigma(d_0)(e_{\mathrm{FF}})$. The considered fake factors are binned in $p_T(e_{\mathrm{FF}})$. The ID and anti-ID criteria are defined using the `Gradient` isolation selection in the electron measurement region with $N_{\text{b-jets}} \geq 1$.



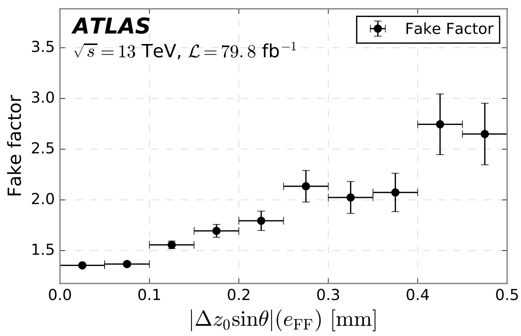
(a) Fake factor as a function of $p_T(e_{\mathrm{FF}})$.



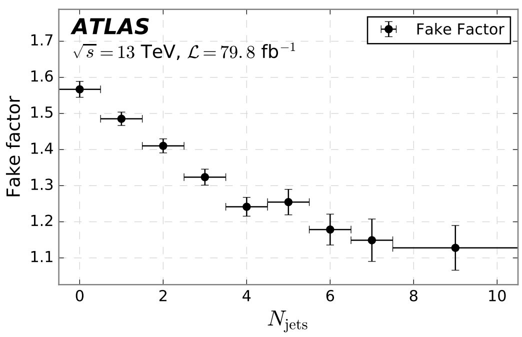
(b) Fake factor as a function of $|\eta|(e_{\mathrm{FF}})$.



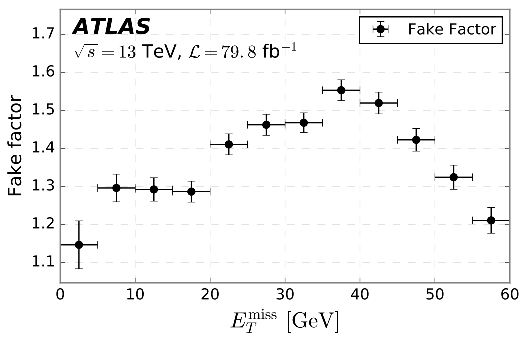
(c) Fake factor as a function of $|d_0|/\sigma(d_0)(e_{\mathrm{FF}})$.



(d) Fake factor as a function of $|\Delta z_0 \sin(\theta)|(e_{\mathrm{FF}})$.

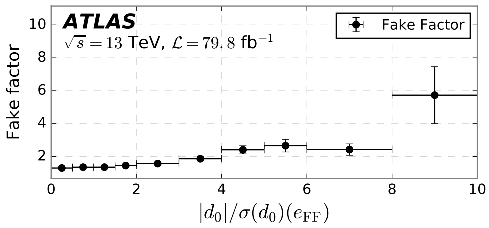


(e) Fake factor as a function of N_{jets}.

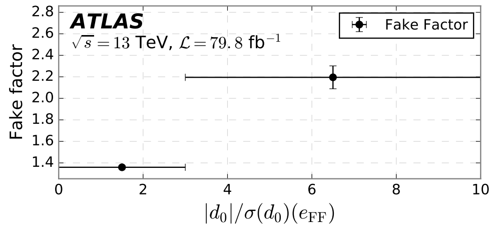


(f) Fake factor as a function of E_T^{miss}.

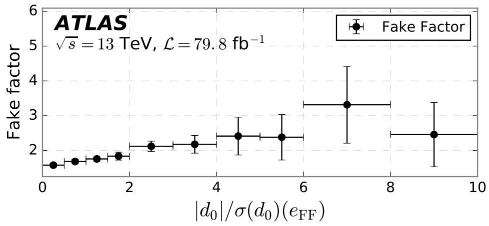
Figure 129: Electron fake factor as a function of: $p_T(e_{\mathrm{FF}})$; $|\eta|(e_{\mathrm{FF}})$; $|d_0|/\sigma(d_0)(e_{\mathrm{FF}})$; $|\Delta z_0 \sin(\theta)|(e_{\mathrm{FF}})$; N_{jets}; and E_T^{miss}. The ID and anti-ID criteria are defined using the Loose isolation in the electron measurement region with $N_{\text{b-jets}} = 0$.



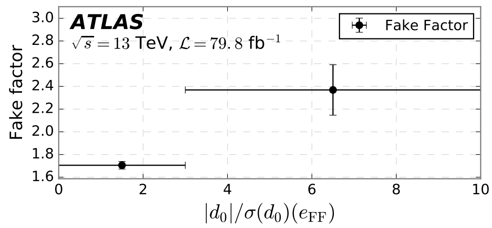
(a) Bin: 10 GeV $\leq p_T(e_{\mathrm{FF}}) < 13$ GeV and $N_{\mathrm{jets}} \leq 2$.



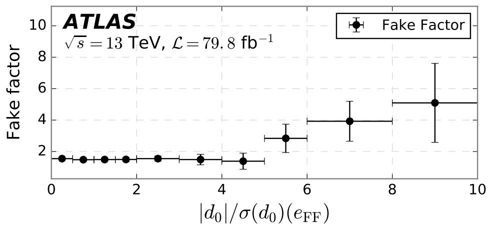
(b) Applied binning for Fig. 130a.



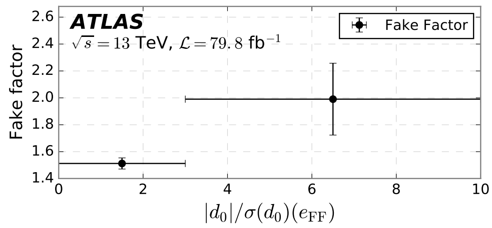
(c) Bin: 13 GeV $\leq p_T(e_{\mathrm{FF}}) < 15$ GeV and $N_{\mathrm{jets}} \leq 2$.



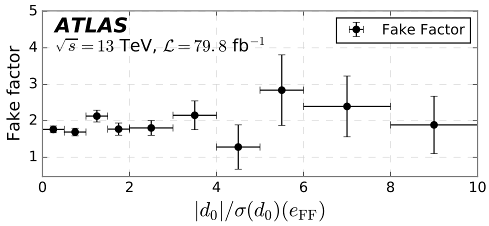
(d) Applied binning for Fig. 130c.



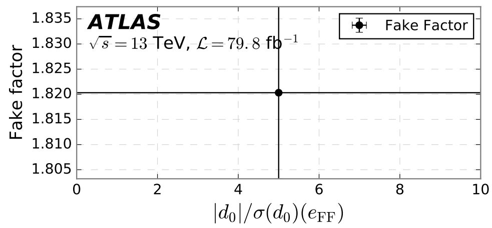
(e) Bin: 15 GeV $\leq p_T(e_{\mathrm{FF}}) < 17$ GeV and $N_{\mathrm{jets}} \leq 2$.



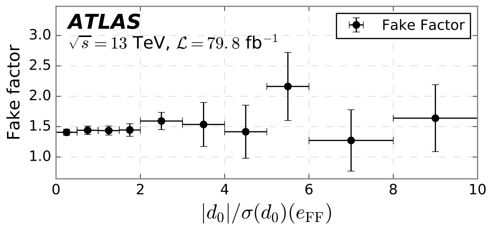
(f) Applied binning for Fig. 130e.



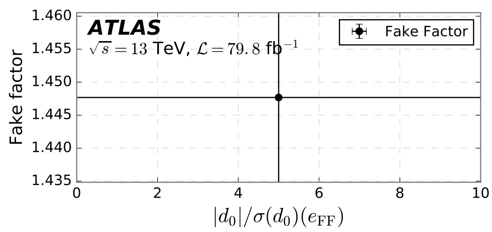
(g) Bin: 17 GeV $\leq p_T(e_{\mathrm{FF}}) < 20$ GeV and $N_{\mathrm{jets}} \leq 2$.



(h) Applied binning for Fig. 130g.

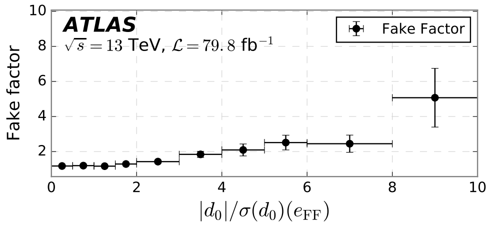


(i) Bin: 20 GeV $\leq p_T(e_{\mathrm{FF}}) < 50$ GeV and $N_{\mathrm{jets}} \leq 2$.

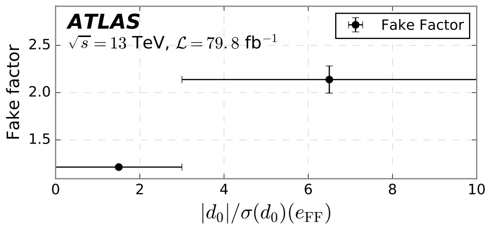


(j) Applied binning for Fig. 130i.

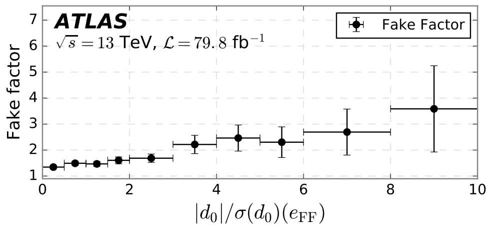
Figure 130: Electron fake factor as a function of $|d_0|/\sigma(d_0)(e_{\mathrm{FF}})$. The considered fake factors are binned in $p_T(\mu_{\mathrm{FF}})$ and N_{jets}. The ID and anti-ID criteria are defined using the `Loose` isolation selection in the electron measurement region with $N_{\text{b-jets}} = 0$.



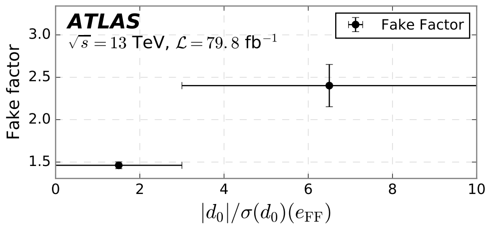
(a) Bin: 10 GeV $\leq p_T(e_{\mathrm{FF}}) <$ 13 GeV and $N_{\mathrm{jets}} > 2$.



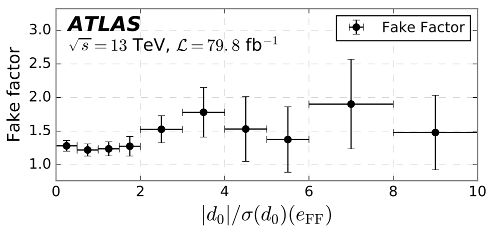
(b) Applied binning for Fig. 131a.



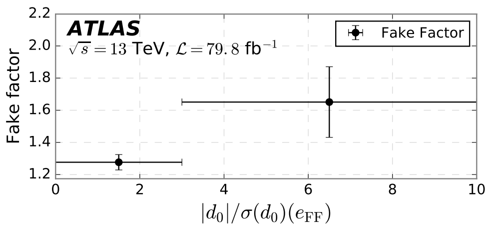
(c) Bin: 13 GeV $\leq p_T(e_{\mathrm{FF}}) <$ 15 GeV and $N_{\mathrm{jets}} > 2$.



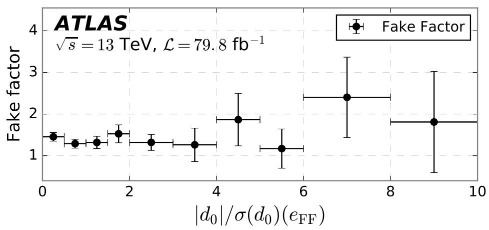
(d) Applied binning for Fig. 131c.



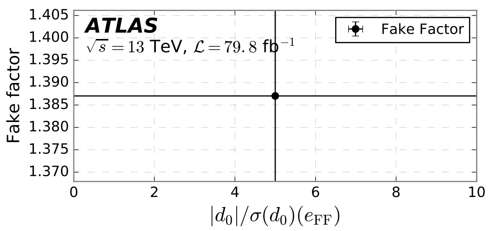
(e) Bin: 15 GeV $\leq p_T(e_{\mathrm{FF}}) <$ 17 GeV and $N_{\mathrm{jets}} > 2$.



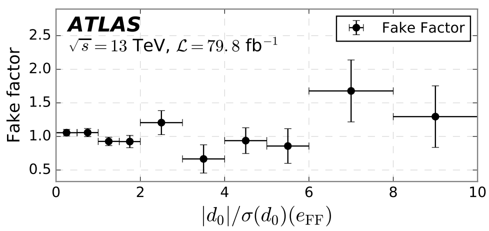
(f) Applied binning for Fig. 131e.



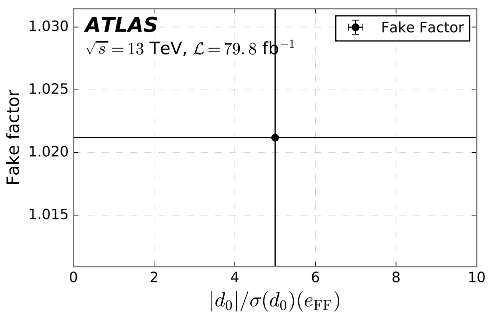
(g) Bin: 17 GeV $\leq p_T(e_{\mathrm{FF}}) <$ 20 GeV and $N_{\mathrm{jets}} > 2$.



(h) Applied binning for Fig. 131g.

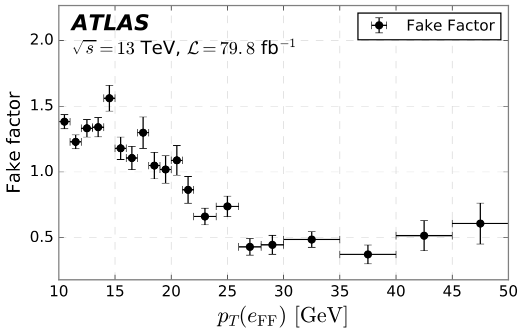


(i) Bin: 20 GeV $\leq p_T(e_{\mathrm{FF}}) <$ 50 GeV and $N_{\mathrm{jets}} > 2$.

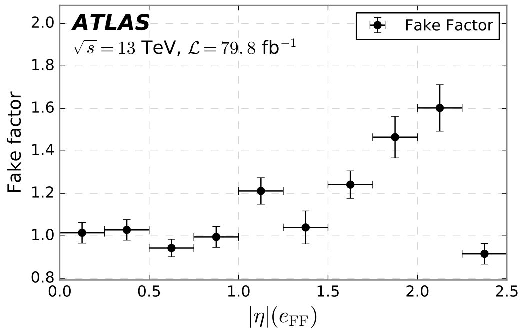


(j) Applied binning for Fig. 131i.

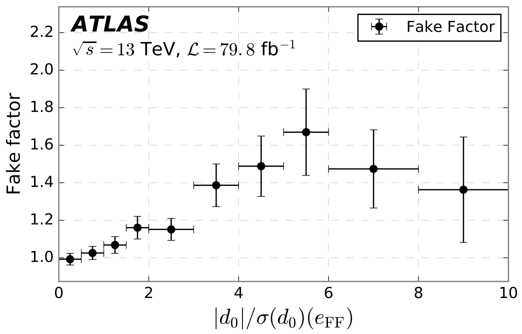
Figure 131: Electron fake factor as a function of $|d_0|/\sigma(d_0)(e_{\mathrm{FF}})$. The considered fake factors are binned in $p_T(\mu_{\mathrm{FF}})$ and N_{jets}. The ID and anti-ID criteria are defined using the `Loose` isolation selection in the electron measurement region with $N_{\text{b-jets}} = 0$.



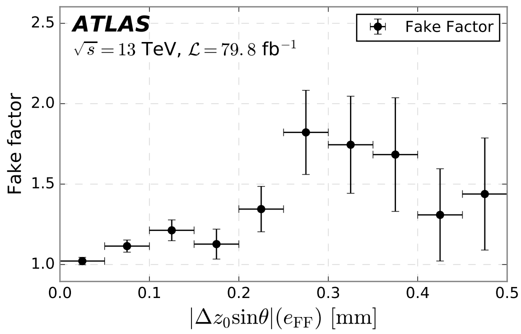
(a) Fake factor as a function of $p_T(e_{\rm FF})$.



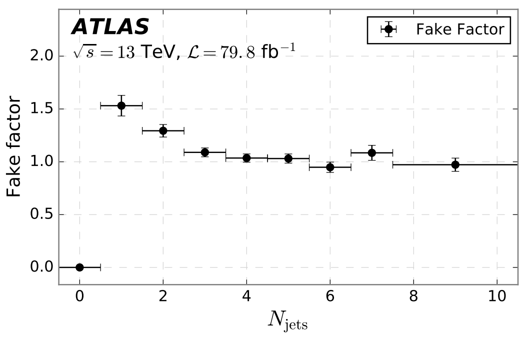
(b) Fake factor as a function of $|\eta|(e_{\rm FF})$.



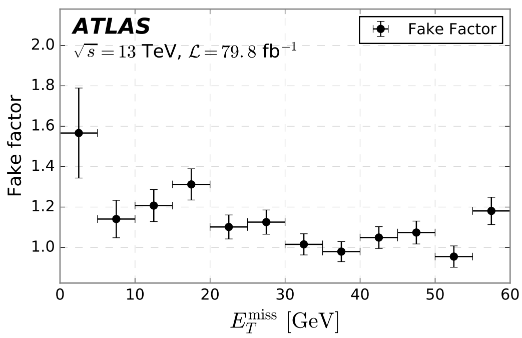
(c) Fake factor as a function of $|d_0|/\sigma(d_0)(e_{\rm FF})$.



(d) Fake factor as a function of $|\Delta z_0 \sin(\theta)|(e_{\rm FF})$.

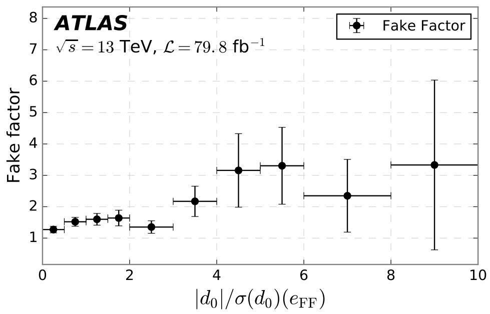


(e) Fake factor as a function of $N_{\rm jets}$.

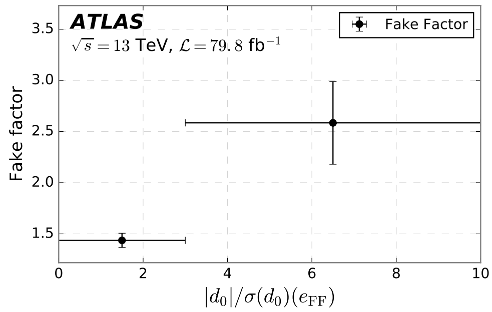


(f) Fake factor as a function of $E_T^{\rm miss}$.

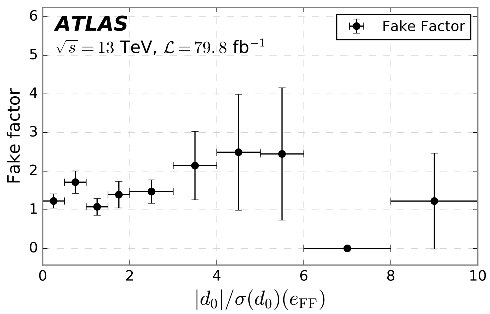
Figure 132: Electron fake factor as a function of: $p_T(e_{\rm FF})$; $|\eta|(e_{\rm FF})$; $|d_0|/\sigma(d_0)(e_{\rm FF})$; $|\Delta z_0 \sin(\theta)|(e_{\rm FF})$; $N_{\rm jets}$; and $E_T^{\rm miss}$. The ID and anti-ID criteria are defined using the `Loose` isolation in the electron measurement region with $N_{\text{b-jets}} \geq 1$.



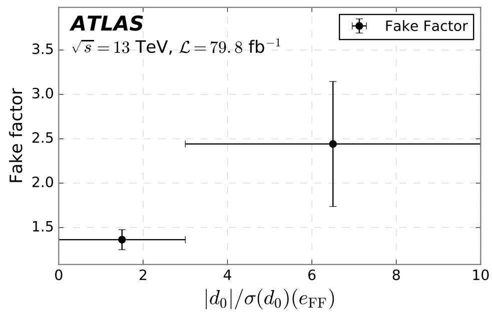
(a) Bin: 10 GeV $\leq p_T(e_{\mathrm{FF}}) < 15$ GeV and $N_{\mathrm{jets}} \leq 2$.



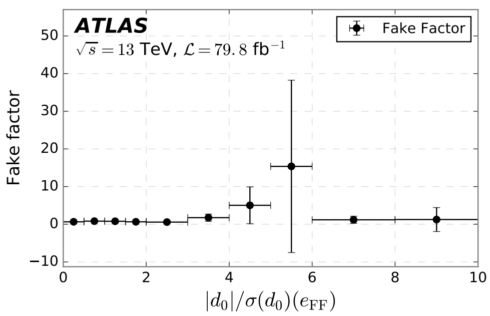
(b) Applied binning for Fig. 133a.



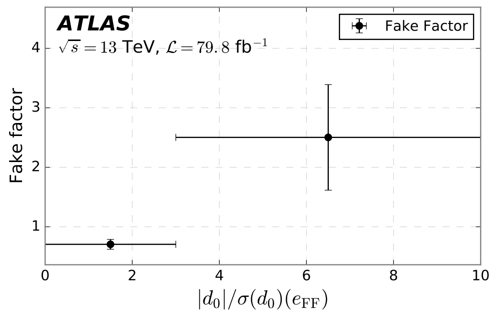
(c) Bin: 15 GeV $\leq p_T(e_{\mathrm{FF}}) < 22$ GeV and $N_{\mathrm{jets}} \leq 2$.



(d) Applied binning for Fig. 133c.

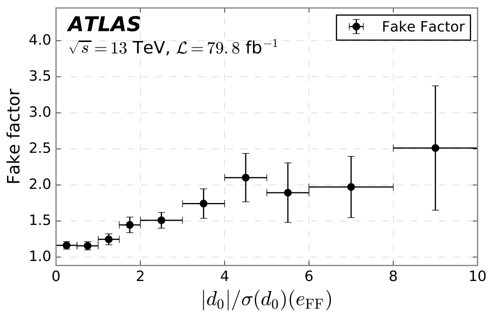


(e) Bin: 22 GeV $\leq p_T(e_{\mathrm{FF}}) < 50$ GeV and $N_{\mathrm{jets}} \leq 2$.

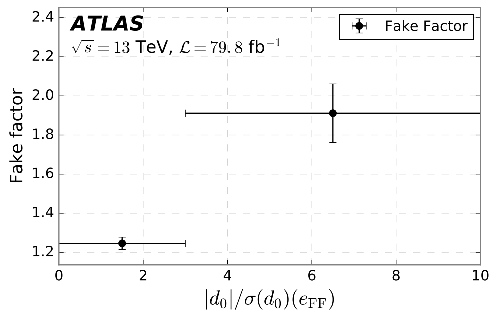


(f) Applied binning for Fig. 133e.

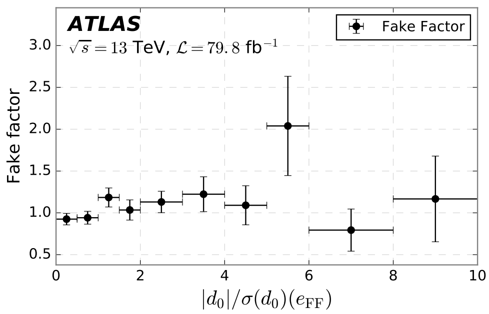
Figure 133: Electron fake factor as a function of $|d_0|/\sigma(d_0)(e_{\mathrm{FF}})$. The considered fake factors are binned in $p_T(\mu_{\mathrm{FF}})$ and N_{jets}. The ID and anti-ID criteria are defined using the `Loose` isolation selection in the electron measurement region with $N_{\text{b-jets}} \geq 1$.



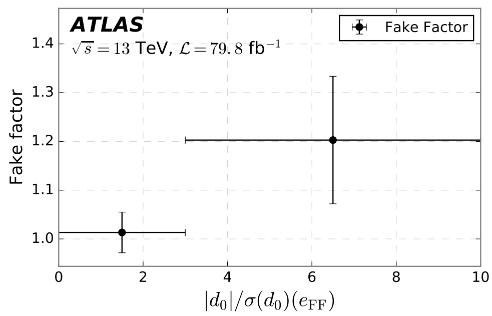
(a) Bin: 10 GeV $\leq p_T(e_{\mathrm{FF}}) < 15$ GeV and $N_{\mathrm{jets}} > 2$.



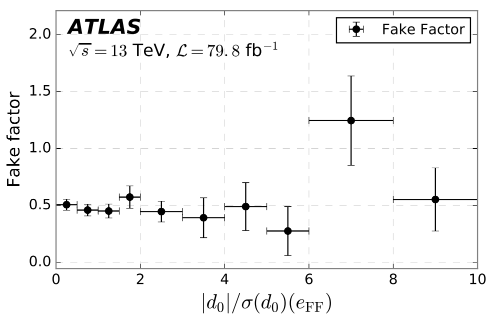
(b) Applied binning for Fig. 134a.



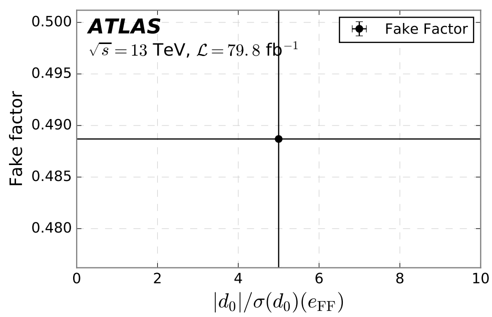
(c) Bin: 15 GeV $\leq p_T(e_{\mathrm{FF}}) < 22$ GeV and $N_{\mathrm{jets}} > 2$.



(d) Applied binning for Fig. 134c.



(e) Bin: 22 GeV $\leq p_T(e_{\mathrm{FF}}) < 50$ GeV and $N_{\mathrm{jets}} > 2$.



(f) Applied binning for Fig. 134e.

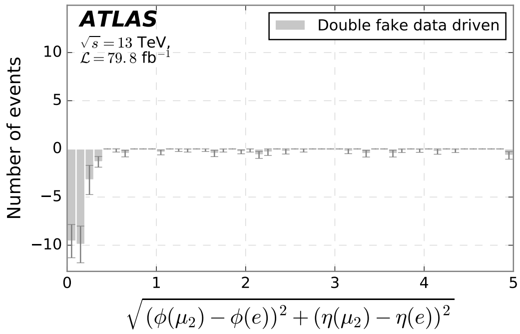
Figure 134: Electron fake factor as a function of $|d_0|/\sigma(d_0)(e_{\mathrm{FF}})$. The considered fake factors are binned in $p_T(\mu_{\mathrm{FF}})$ and N_{jets}. The ID and anti-ID criteria are defined using the `Loose` isolation selection in the electron measurement region with $N_{\text{b-jets}} \geq 1$.

G Multi-fake background estimation

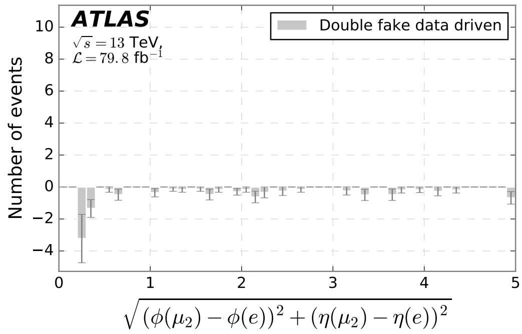
When applying the fake factor method in a possible multi-fake region, it is assumed that the fake leptons are independent of each other. This is a good assumption when the trajectories of the two fake leptons are separated but not necessarily in the cases where the fake lepton tracks are close to each.

When both the subleading muon and electron are fake leptons and have a small angular separation, they can both origin from the same jet. The overlap removal between the reconstructed muon and electron object, see Tab. 6, is only applied if the two objects share the same tracks without considering the ΔR separation of the tracks. In comparison, the overlap removal between a reconstructed jet and electron or muon will in general remove the jet if $\Delta R(\text{jet}, \mu/e) < 0.2$.

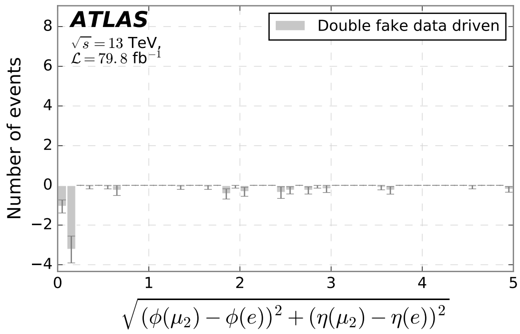
Figure 135 shows the data-driven estimate of the double fake background, see Eq. 8.8, for the $\Delta R(\mu_2, e)$ distribution in control region 1, 2 and 3. The two fake leptons are clearly not independent when the $\Delta R(\mu_2, e) > 0.2$ requirement is not applied. When applying the $\Delta R(\mu_2, e) > 0.2$ requirement, it is ensured that the fake leptons are more independent. Fig. 135b shows that there is still larger negative bins at small $\Delta R(\mu_2, e)$ for control region 1. Therefore, it can be necessary to increase the applied $\Delta R(\mu_2, e)$ requirement in order to ensure that the leptons are truly independent. The requirement $\Delta R(\mu_2, e)$ is chosen such that the efficiency for the signal sample with $M_N = 5$ GeV is not reduced much, see Fig. 21a.



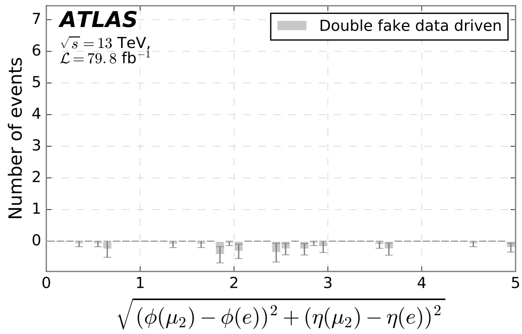
(a) Control region 1:
No $\Delta R(\mu_2, 3) > 0.2$ requirement applied.



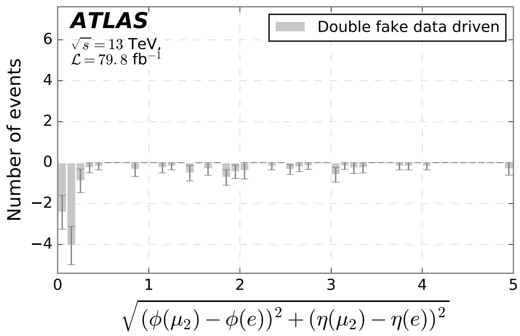
(b) Control region 1:
With $\Delta R(\mu_2, 3) > 0.2$ requirement applied.



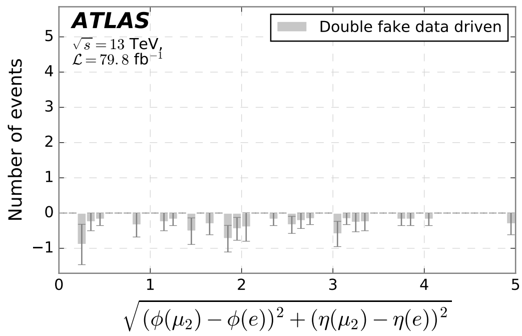
(c) Control region 2:
No $\Delta R(\mu_2, 3) > 0.2$ requirement applied.



(d) Control region 2:
With $\Delta R(\mu_2, 3) > 0.2$ requirement applied.



(e) Control region 3:
No $\Delta R(\mu_2, 3) > 0.2$ requirement applied.



(f) Control region 3:
With $\Delta R(\mu_2, 3) > 0.2$ requirement applied.

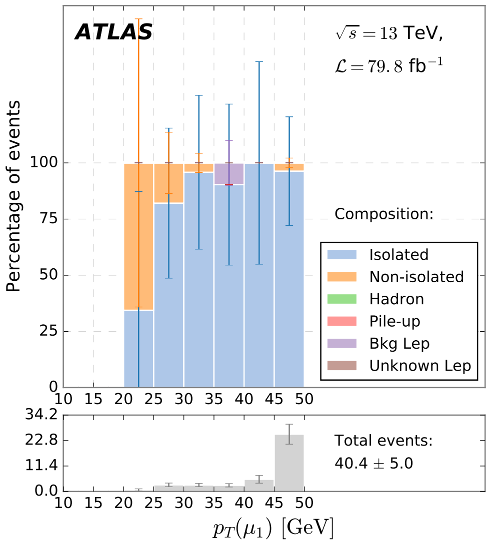
Figure 135: Control region 1, 2 and 3: $\Delta R(\mu_2, e)$ distribution. Only the data-driven estimate of the double fake background, which has a negative sign, is shown, see Eq. 8.8. Isolation for μ_2: `FixedCutLoose`. Isolation for e: `Gradient`.

H Composition study of control regions

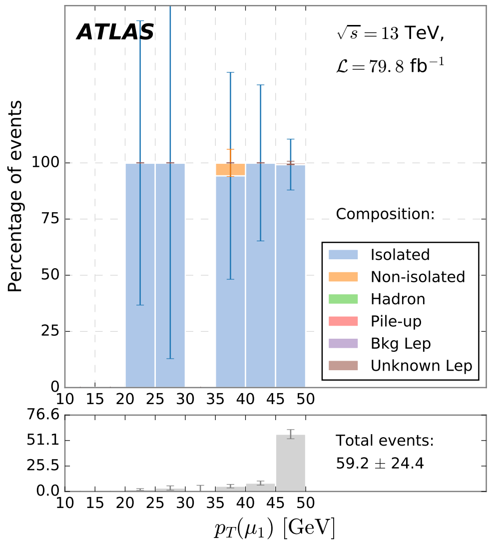
Three control regions are applied in order to examine the agreement between data and the background estimation. The control regions are defined kinematically close to the signal region. In this section, the composition of the final state leptons in the control regions is studied using truth matching. The figures applied in the composition study and the main results are presented in Tab. 75.

Composition	**Figures**	**Result**
Leading muon	Fig. 136-138	Large majority are prompt muons. Few fake muons of type "non-isolated".
Subleading muon	Fig. 139-141	Large majority are fake muons of type "non-isolated". Contribution from "pile-up" muons in the control region with anti-ID criteria for both μ_2 and e.
Electron	Fig. 142-144	Majority of fake electrons of type "non-isolated" and "hadron". The primary contribution changes for the three control regions. Control region 2 is dominated by "non-isolated" fake electrons. Control region 1 and 3 have larger contribution from "hadron".

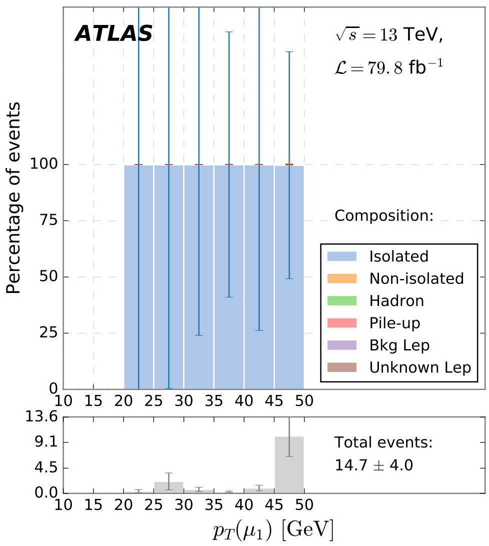
Table 75: Summary of the composition study for the three control regions. Control region 1: Inverted $M(\mu_1, \mu_2, e)$ requirement. Control region 2: Inverted b-jet veto. Control region 3: Inverted E_T^{miss} requirement.



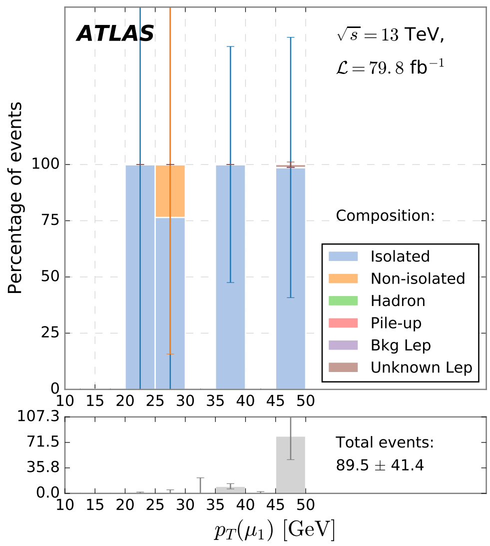
(a) Control region 1. Composition of μ_1. Region: μ_2 : ID criteria. e : ID criteria.



(b) Control region 1. Composition of μ_1. Region: μ_2 : Anti-ID criteria. e : ID criteria.

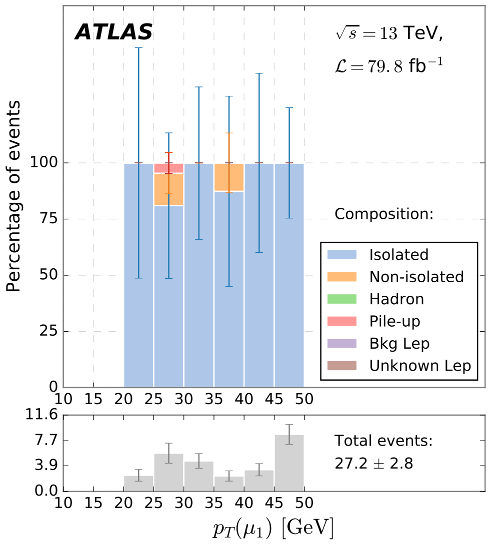


(c) Control region 1. Composition of μ_1. Region: μ_2 : ID criteria. e : Anti-ID criteria.

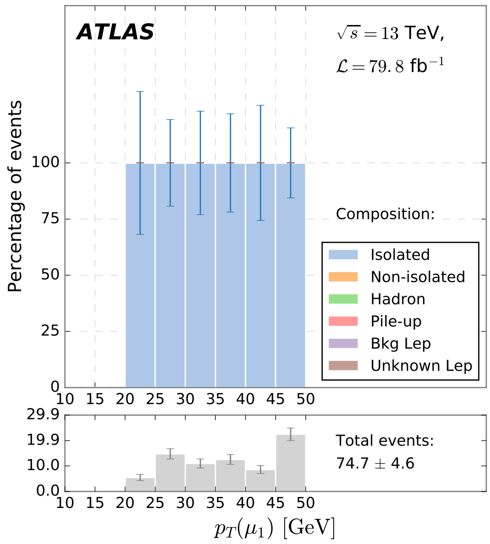


(d) Control region 1. Composition of μ_1. Region: μ_2 : Anti-ID criteria. e : Anti-ID criteria.

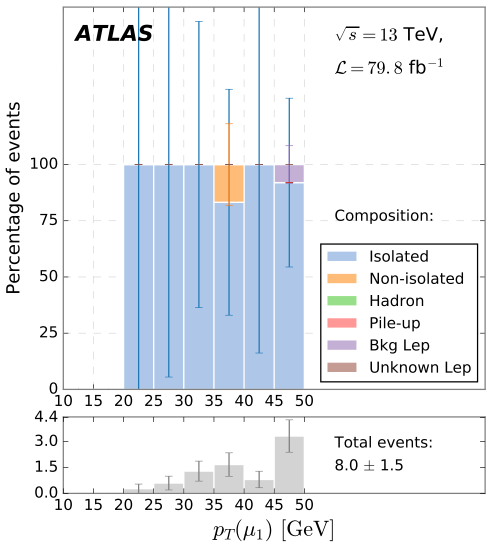
Figure 136: Composition of μ_1 in control region 1 as a function of $p_T(\mu_1)$. The final bin contains all the events with $p_T > 45$ GeV. The main plot shows the composition as the percentage of events in the p_T bin. In the subplot, the number of events in each p_T bin is shown as well as the total number of events.



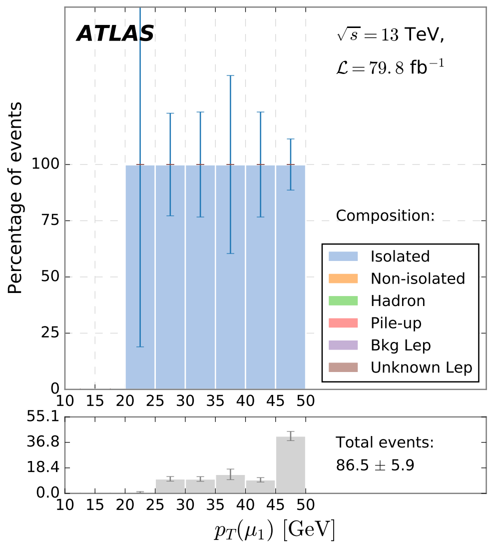
(a) Control region 2. Composition of μ_1. Region: μ_2 : ID criteria. e : ID criteria.



(b) Control region 2. Composition of μ_1. Region: μ_2 : Anti-ID criteria. e : ID criteria.

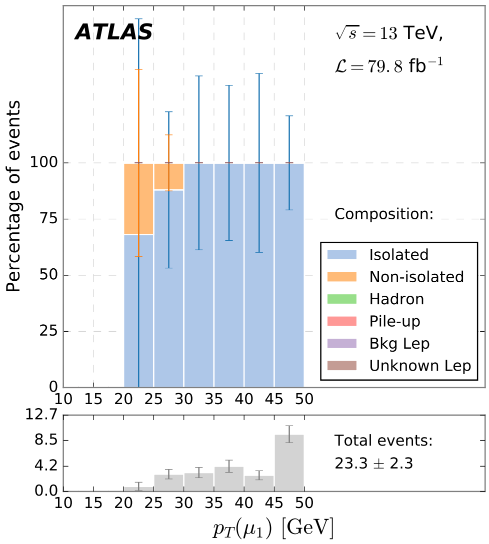


(c) Control region 2. Composition of μ_1. Region: μ_2 : ID criteria. e : Anti-ID criteria.

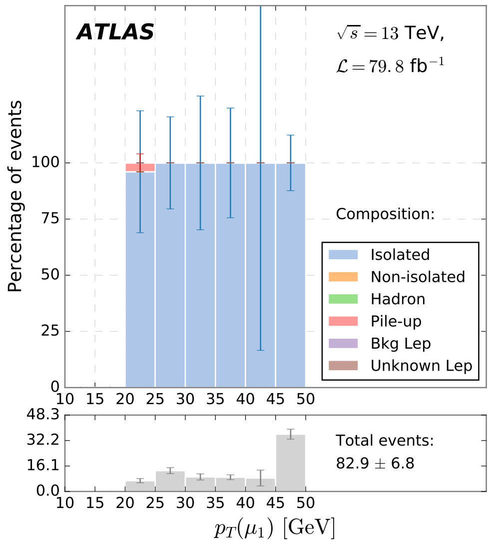


(d) Control region 2. Composition of μ_1. Region: μ_2 : Anti-ID criteria. e : Anti-ID criteria.

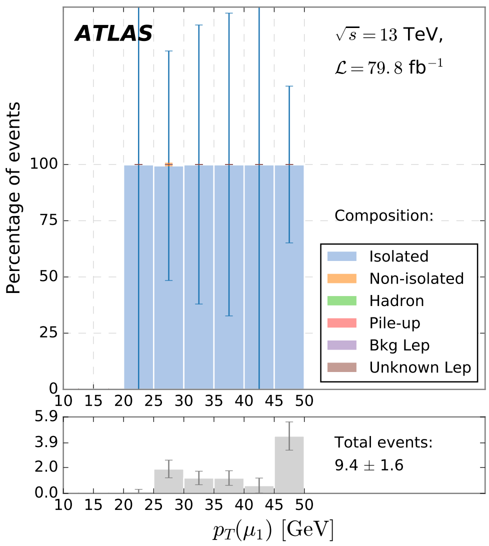
Figure 137: Composition of μ_1 in control region 2 as a function of $p_T(\mu_1)$. The final bin contains all the events with $p_T > 45$ GeV. The main plot shows the composition as the percentage of events in the p_T bin. In the subplot, the number of events in each p_T bin is shown as well as the total number of events.



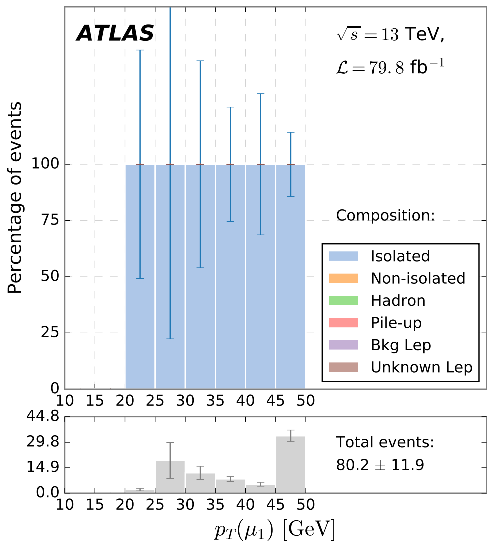
(a) Control region 3. Composition of μ_1. Region: μ_2 : ID criteria. e : ID criteria.



(b) Control region 3. Composition of μ_1. Region: μ_2 : Anti-ID criteria. e : ID criteria.

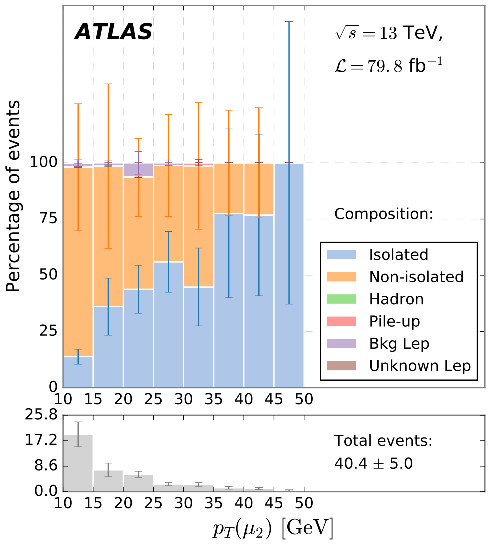


(c) Control region 3. Composition of μ_1. Region: μ_2 : ID criteria. e : Anti-ID criteria.

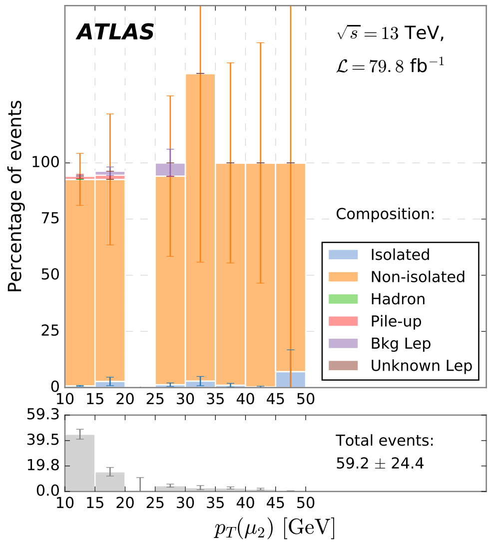


(d) Control region 3. Composition of μ_1. Region: μ_2 : Anti-ID criteria. e : Anti-ID criteria.

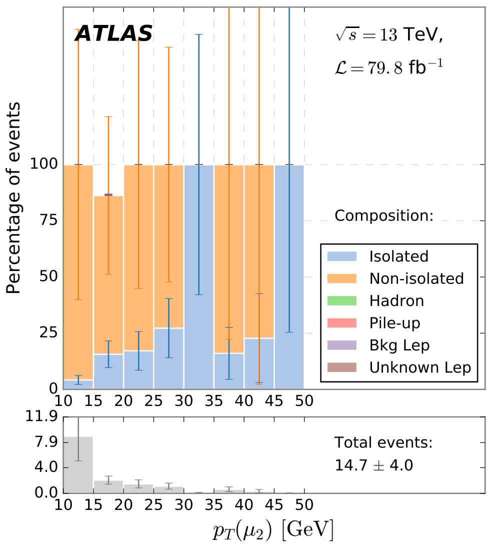
Figure 138: Composition of μ_1 in control region 3 as a function of $p_T(\mu_1)$. The final bin contains all the events with $p_T > 45$ GeV. The main plot shows the composition as the percentage of events in the p_T bin. In the subplot, the number of events in each p_T bin is shown as well as the total number of events.



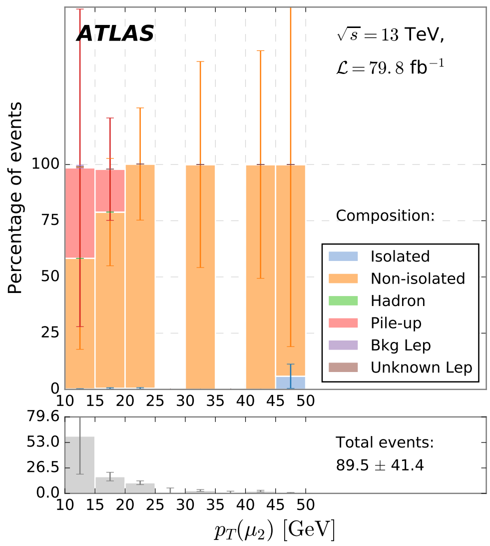
(a) Control region 1. Composition of μ_1.
Region: μ_2 : ID criteria. e : ID criteria.



(b) Control region 1. Composition of μ_1.
Region: μ_2 : Anti-ID criteria. e : ID criteria.

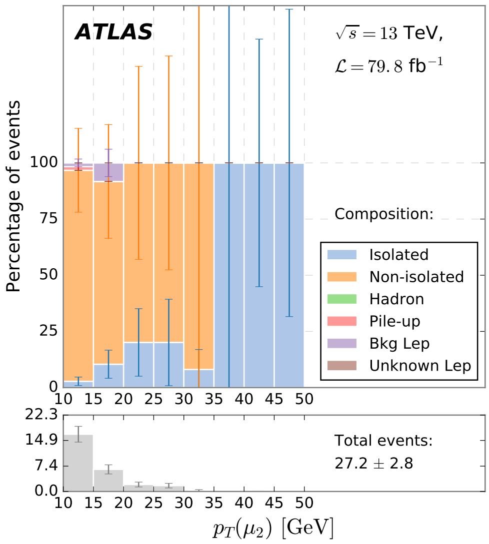


(c) Control region 1. Composition of μ_1.
Region: μ_2 : ID criteria. e : Anti-ID criteria.

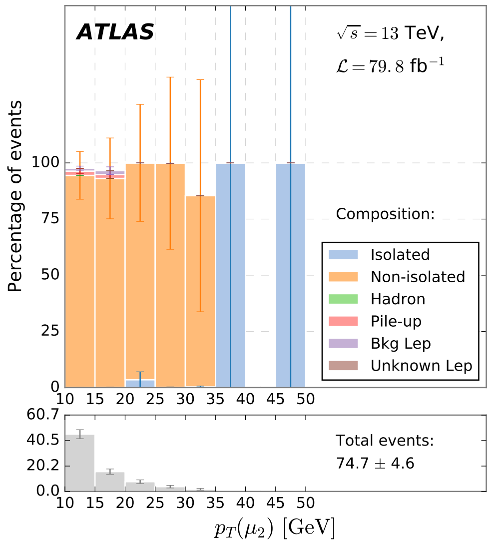


(d) Control region 1. Composition of μ_1.
Region: μ_2 : Anti-ID criteria. e : Anti-ID criteria.

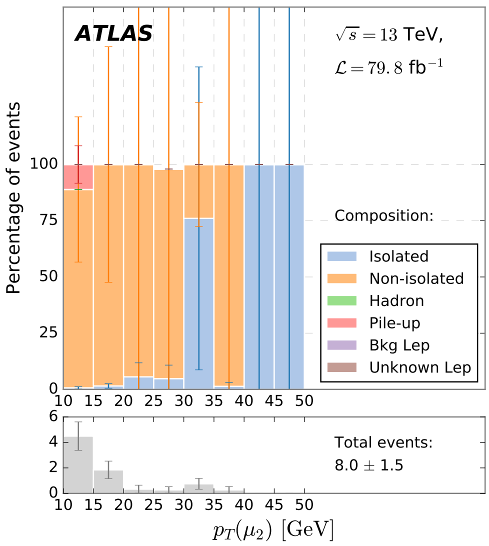
Figure 139: Composition of μ_2 in control region 1 as a function of $p_T(\mu_2)$. The final bin contains all the events with $p_T > 45$ GeV. The main plot shows the composition as the percentage of events in the p_T bin. In the subplot, the number of events in each p_T bin is shown as well as the total number of events.



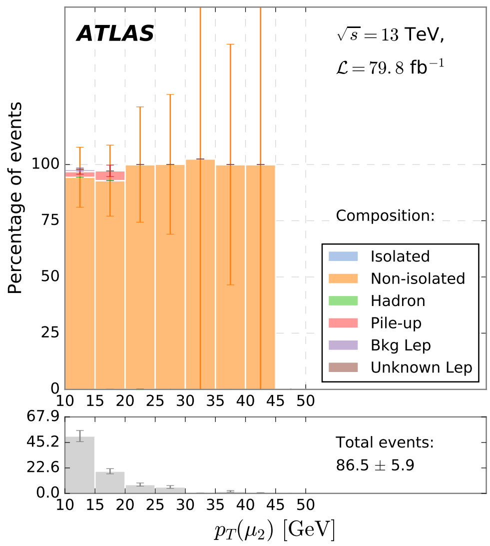
(a) Control region 2. Composition of μ_1.
Region: μ_2 : ID criteria. e : ID criteria.



(b) Control region 2. Composition of μ_1.
Region: μ_2 : Anti-ID criteria. e : ID criteria.

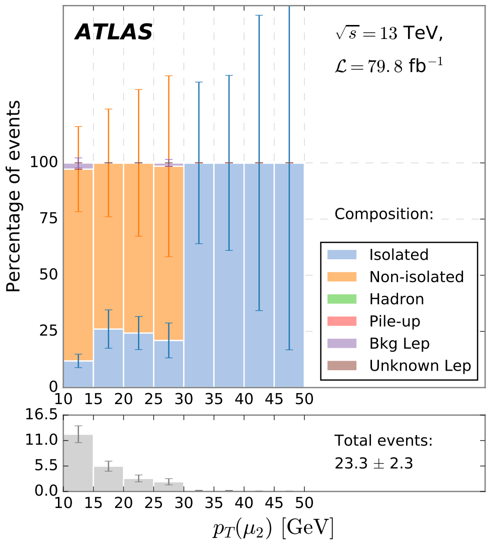


(c) Control region 2. Composition of μ_1.
Region: μ_2 : ID criteria. e : Anti-ID criteria.

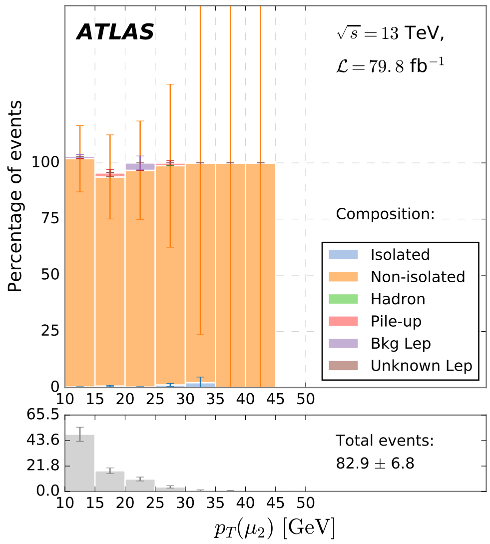


(d) Control region 2. Composition of μ_1.
Region: μ_2 : Anti-ID criteria. e : Anti-ID criteria.

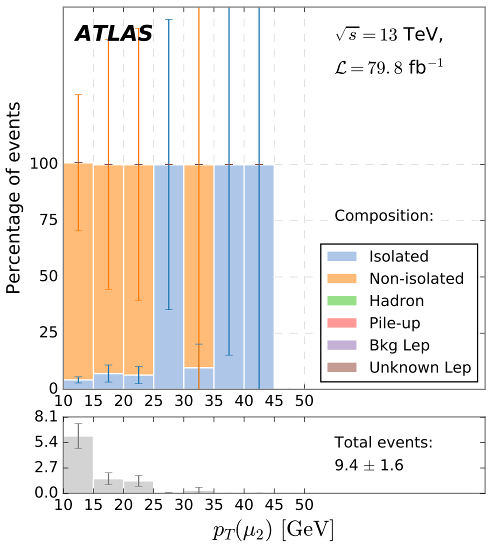
Figure 140: Composition of μ_2 in control region 2 as a function of $p_T(\mu_2)$. The final bin contains all the events with $p_T > 45$ GeV. The main plot shows the composition as the percentage of events in the p_T bin. In the subplot, the number of events in each p_T bin is shown as well as the total number of events.



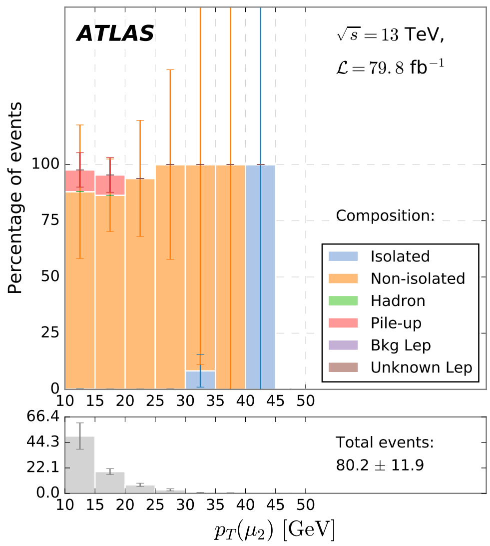
(a) Control region 3. Composition of μ_1. Region: μ_2 : ID criteria. e : ID criteria.



(b) Control region 3. Composition of μ_1. Region: μ_2 : Anti-ID criteria. e : ID criteria.

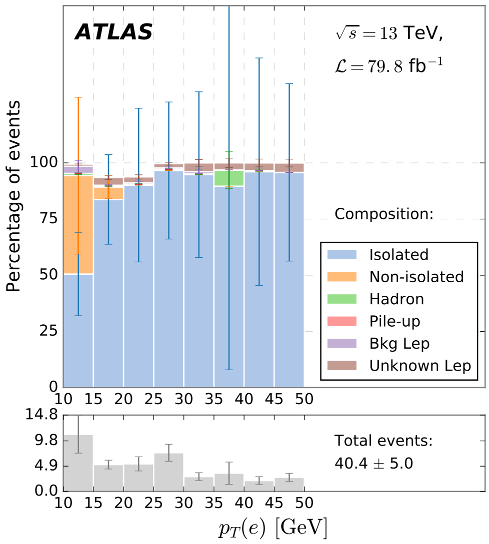


(c) Control region 3. Composition of μ_1. Region: μ_2 : ID criteria. e : Anti-ID criteria.

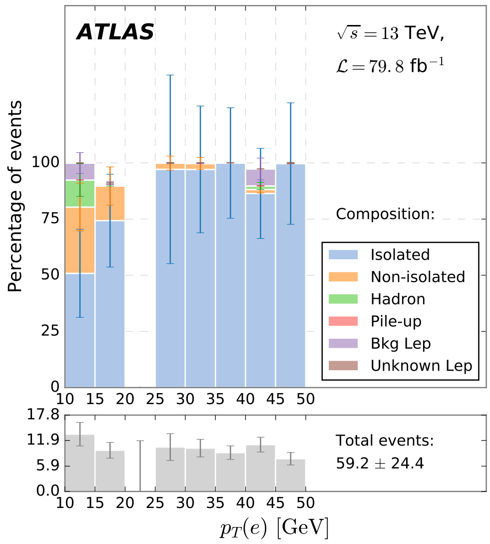


(d) Control region 3. Composition of μ_1. Region: μ_2 : Anti-ID criteria. e : Anti-ID criteria.

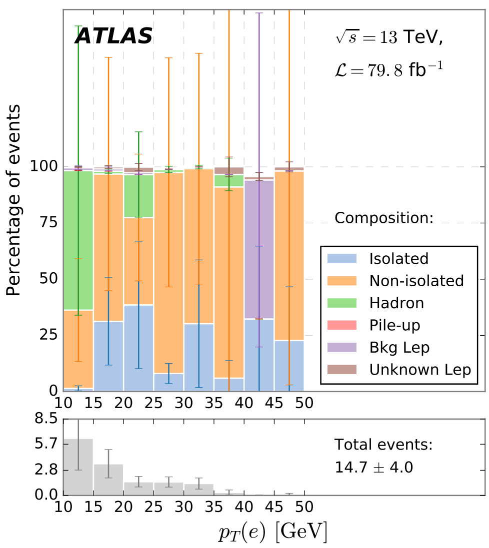
Figure 141: Composition of μ_2 in control region 3 as a function of $p_T(\mu_2)$. The final bin contains all the events with $p_T > 45$ GeV. The main plot shows the composition as the percentage of events in the p_T bin. In the subplot, the number of events in each p_T bin is shown as well as the total number of events.



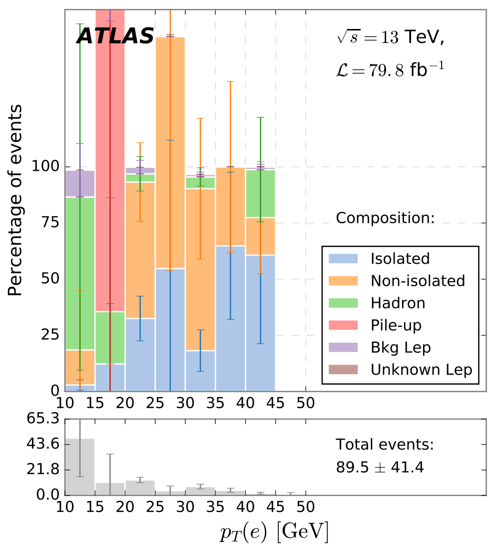
(a) Control region 1. Composition of μ_1.
Region: μ_2 : ID criteria. e : ID criteria.



(b) Control region 1. Composition of μ_1.
Region: μ_2 : Anti-ID criteria. e : ID criteria.

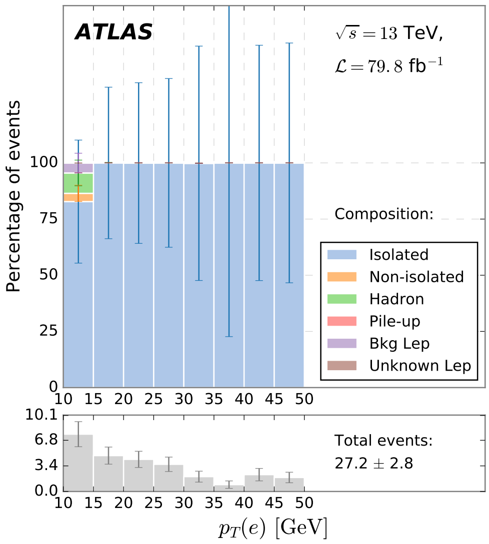


(c) Control region 1. Composition of μ_1.
Region: μ_2 : ID criteria. e : Anti-ID criteria.

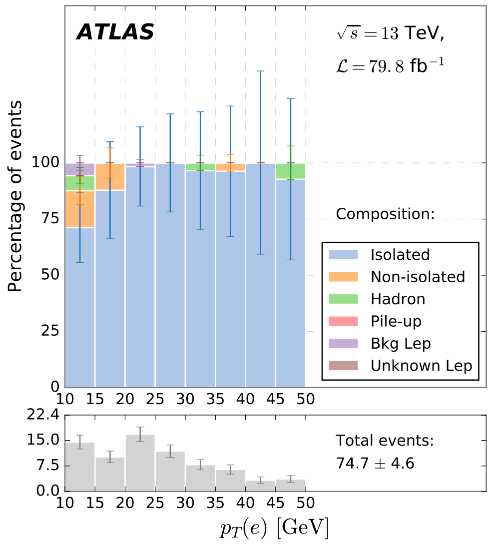


(d) Control region 1. Composition of μ_1.
Region: μ_2 : Anti-ID criteria. e : Anti-ID criteria.

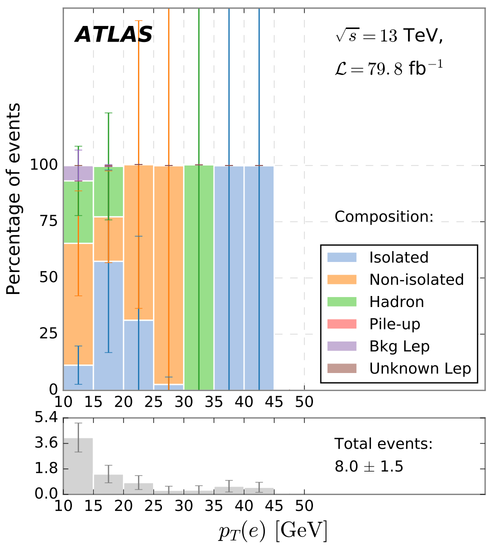
Figure 142: Composition of e in control region 1 as a function of $p_T(e)$. The final bin contains all the events with $p_T > 45$ GeV. The main plot shows the composition as the percentage of events in the p_T bin. In the subplot, the number of events in each p_T bin is shown as well as the total number of events.



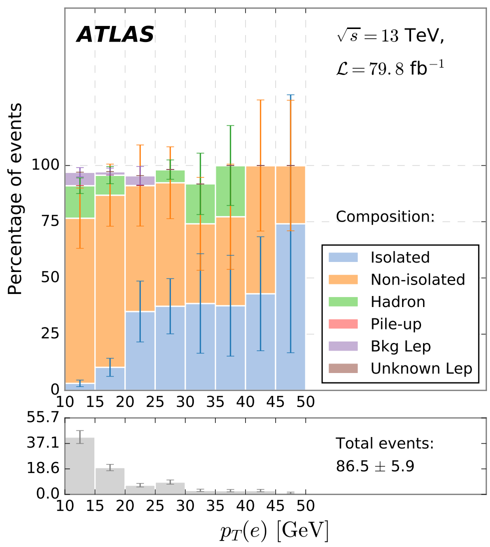
(a) Control region 2. Composition of μ_1. Region: μ_2 : ID criteria. e : ID criteria.



(b) Control region 2. Composition of μ_1. Region: μ_2 : Anti-ID criteria. e : ID criteria.

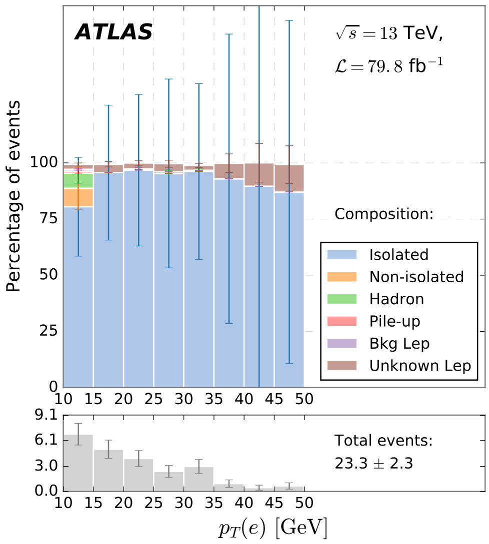


(c) Control region 2. Composition of μ_1. Region: μ_2 : ID criteria. e : Anti-ID criteria.

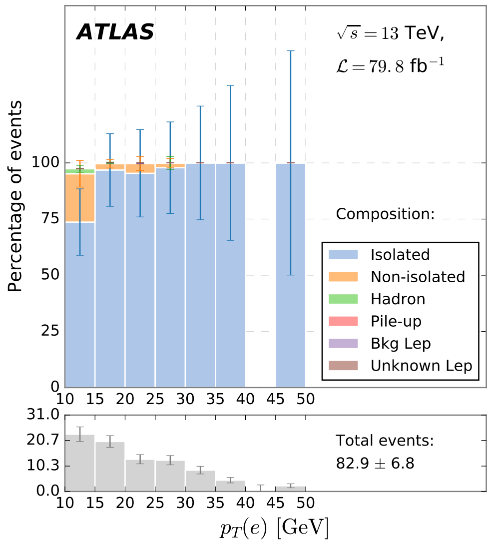


(d) Control region 2. Composition of μ_1. Region: μ_2 : Anti-ID criteria. e : Anti-ID criteria.

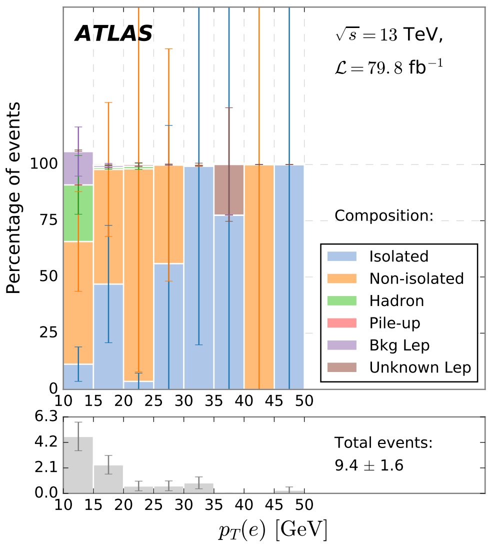
Figure 143: Composition of e in control region 2 as a function of $p_T(e)$. The final bin contains all the events with $p_T > 45$ GeV. The main plot shows the composition as the percentage of events in the p_T bin. In the subplot, the number of events in each p_T bin is shown as well as the total number of events.



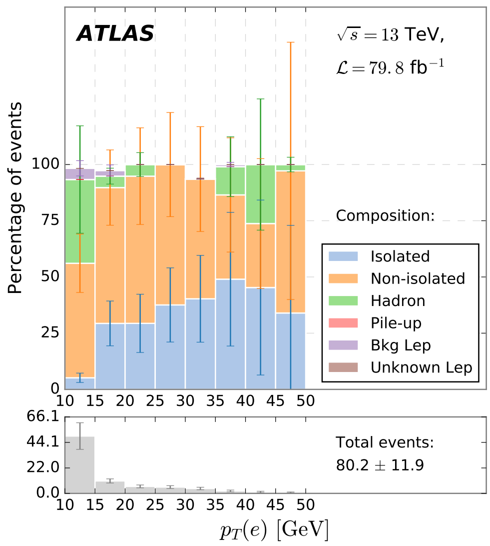
(a) Control region 3. Composition of μ_1. Region: μ_2 : ID criteria. e : ID criteria.



(b) Control region 3. Composition of μ_1. Region: μ_2 : Anti-ID criteria. e : ID criteria.



(c) Control region 3. Composition of μ_1. Region: μ_2 : ID criteria. e : Anti-ID criteria.



(d) Control region 3. Composition of μ_1. Region: μ_2 : Anti-ID criteria. e : Anti-ID criteria.

Figure 144: Composition of e in control region 3 as a function of $p_T(e)$. The final bin contains all the events with $p_T > 45$ GeV. The main plot shows the composition as the percentage of events in the p_T bin. In the subplot, the number of events in each p_T bin is shown as well as the total number of events.

I Studies used in the systematic uncertainty estimates

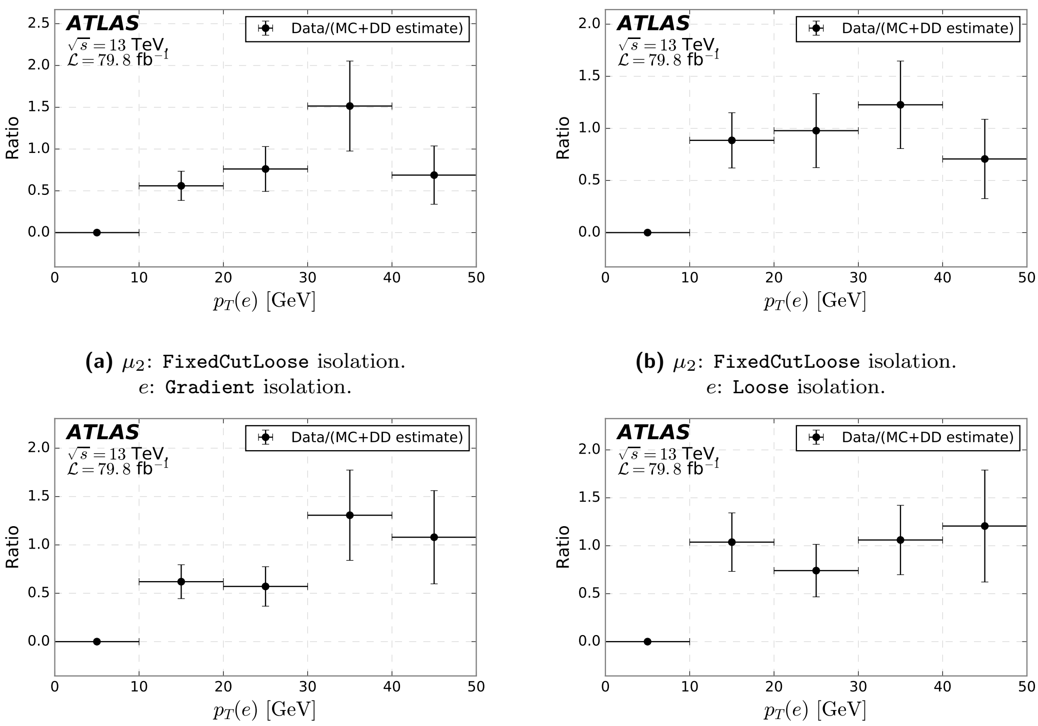
In this section, studies are presented which are used to assign systematic uncertainties to the fake factors applied in the data-driven model.

I.1 Ratio of data and estimated background in control region 1, 2 and 3

The data-driven background model is applied in control regions independent of the signal region. The ratio between the number of data events and the estimated background is shown for different distributions. Table 76 lists the figures included in this section.

Figure	Control region	Distribution
Fig. 43	1	$p_T(\mu_2)$
Fig. 145	1	$p_{\mathrm{T}}(e)$
Fig. 146	2	$p_T(\mu_2)$
Fig. 147	2	$p_{\mathrm{T}}(e)$
Fig. 148	3	$p_T(\mu_2)$
Fig. 149	3	$p_{\mathrm{T}}(e)$

Table 76: List of the figures presented in this section.



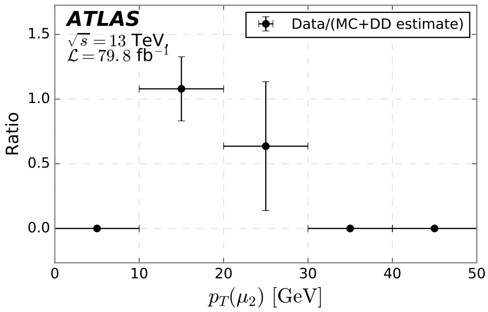
(a) μ_2: `FixedCutLoose` isolation.
e: `Gradient` isolation.

(b) μ_2: `FixedCutLoose` isolation.
e: `Loose` isolation.

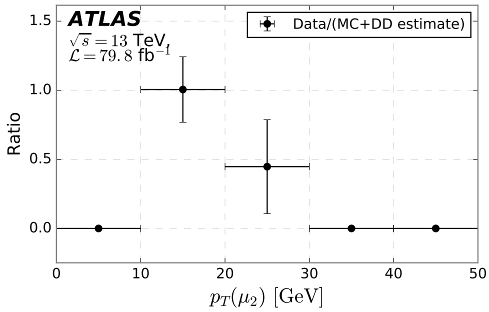
(c) μ_2: `Loosened FixedCutLoose` isolation.
e: `Gradient` isolation.

(d) μ_2: `Loosened FixedCutLoose` isolation.
e: `Loose` isolation.

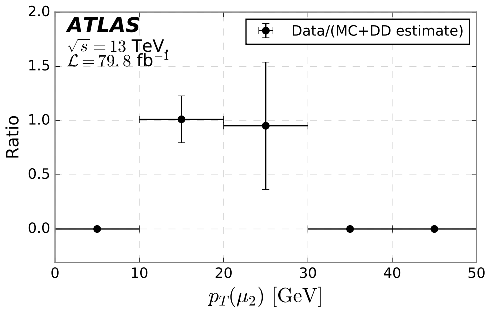
Figure 145: Control Region 1 - Inverted $M(\mu_1, \mu_2, e)$ requirement. Distribution showing $p_{\mathrm{T}}(e)$.



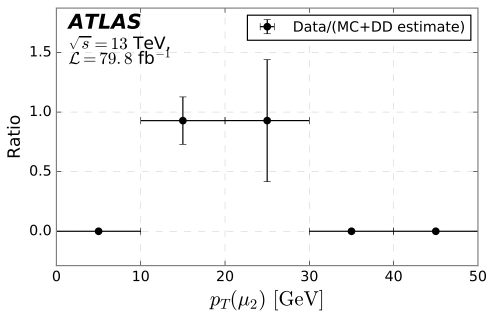
(a) μ_2: `FixedCutLoose` isolation.
e: `Gradient` isolation.



(b) μ_2: `FixedCutLoose` isolation.
e: `Loose` isolation.

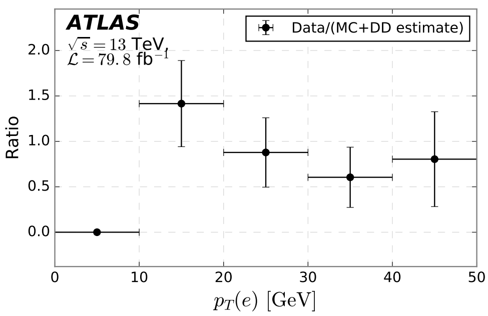


(c) μ_2: `Loosened FixedCutLoose` isolation.
e: `Gradient` isolation.

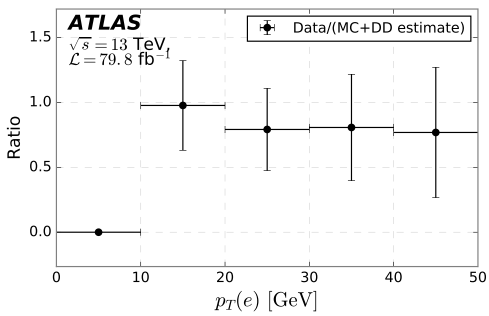


(d) μ_2: `Loosened FixedCutLoose` isolation.
e: `Loose` isolation.

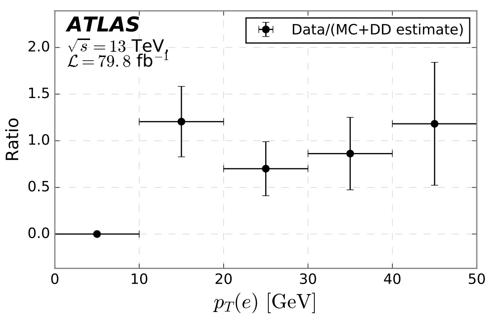
Figure 146: Control Region 2 - Inverted b-jet veto. Distribution showing $p_T(\mu_2)$.



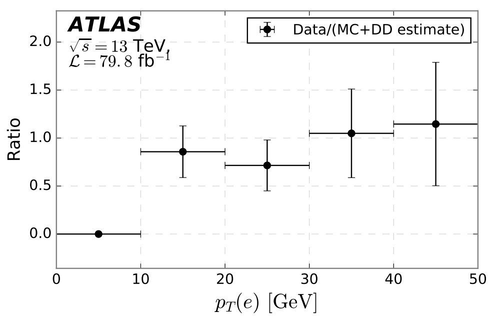
(a) μ_2: `FixedCutLoose` isolation.
e: `Gradient` isolation.



(b) μ_2: `FixedCutLoose` isolation.
e: `Loose` isolation.

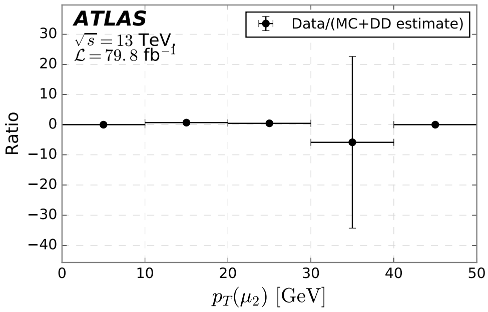


(c) μ_2: `Loosened FixedCutLoose` isolation.
e: `Gradient` isolation.

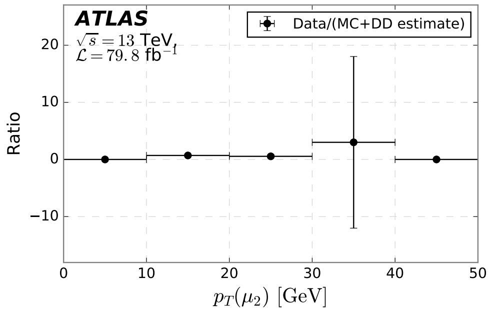


(d) μ_2: `Loosened FixedCutLoose` isolation.
e: `Loose` isolation.

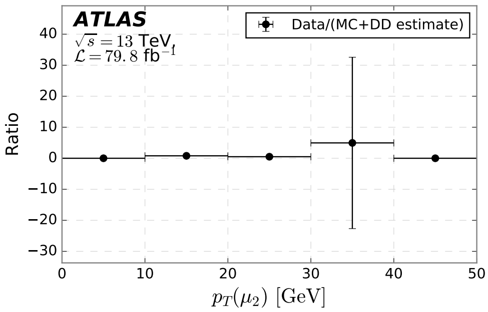
Figure 147: Control Region 2 - Inverted b-jet veto. Distribution showing $p_{\mathrm{T}}(e)$.



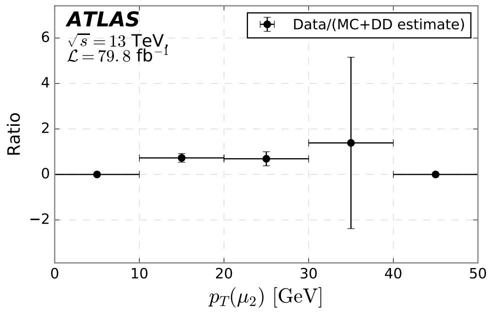
(a) μ_2: `FixedCutLoose` isolation.
e: `Gradient` isolation.



(b) μ_2: `FixedCutLoose` isolation.
e: `Loose` isolation.

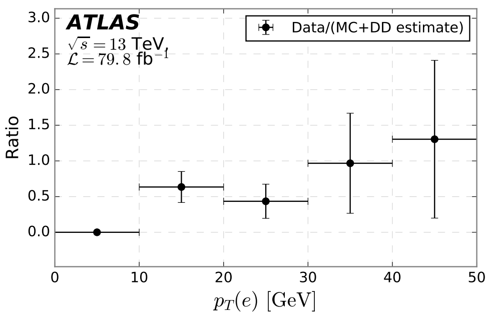


(c) μ_2: `Loosened FixedCutLoose` isolation.
e: `Gradient` isolation.

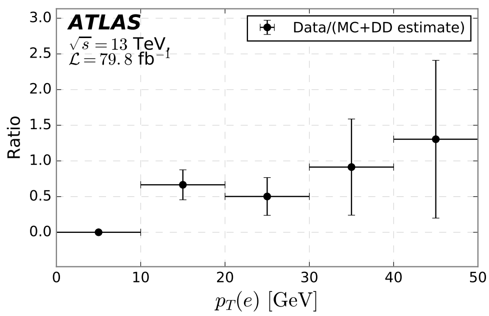


(d) μ_2: `Loosened FixedCutLoose` isolation.
e: `Loose` isolation.

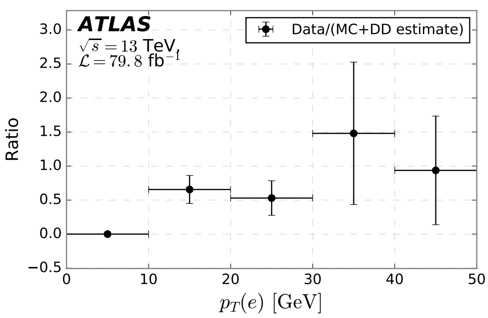
Figure 148: Control Region 2 - Inverted b-jet veto. Distribution showing $p_T(\mu_2)$.



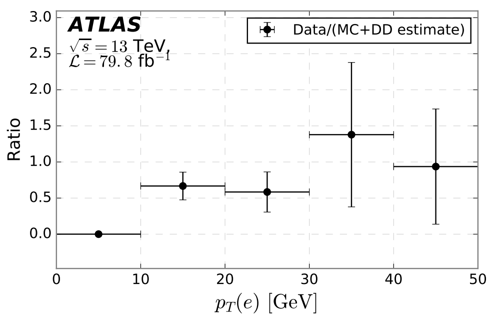
(a) μ_2: `FixedCutLoose` isolation.
e: `Gradient` isolation.



(b) μ_2: `FixedCutLoose` isolation.
e: `Loose` isolation.



(c) μ_2: `Loosened FixedCutLoose` isolation.
e: `Gradient` isolation.



(d) μ_2: `Loosened FixedCutLoose` isolation.
e: `Loose` isolation.

Figure 149: Control Region 3 - Inverted E_T^{miss} requirement. Distribution showing $p_{\text{T}}(e)$.

I.2 MC subtraction

The systematic uncertainty of the fake factors introduced by the MC subtraction is estimated in Sec. 9. The result of this study is shown below. Table 77 shows the percentage difference between the fake factors with and without a 15% scaling of the diboson samples.

			Muon fake factor		Electron fake factor	
p_T bin	$\lvert d_0\rvert/\sigma(d_0)$ bin	N_{jets} bin	Diboson MC +15%	Diboson MC -15%	Diboson MC +15%	Diboson MC -15%
1	1	1	0.7±17.1	-0.7±17.3	0.1±14.0	-0.1±14.0
2	1	1	2.8±22.4	-2.8±23.7	0.2±17.4	-0.2±17.5
3	1	1	6.2±33.6	-6.2±37.1	0.4±16.5	-0.4±16.6
4	1	1	9.1±33.1	-9.0±37.0	1.0±12.4	-1.0±12.6
5	1	1	–	–	3.2±15.2	-3.2±16.2
1	2	1	0.5±14.9	-0.5±15.1	0.1±17.8	-0.1±17.8
2	2	1	4.0±32.7	-4.0±35.4	0.3±22.9	-0.3±23.0
3	2	1	9.8±41.6	-9.7±50.0	0.5±18.2	-0.5±18.4
4	2	1	9.6±28.2	-9.5±33.2	1.6±15.2	-1.6±15.7
5	2	1	–	–	5.1±17.8	-5.0±19.7
1	1	2	0.2±19.2	-0.2±19.3	0.1±24.5	-0.1±24.6
2	1	2	0.8±35.2	-0.8±35.8	0.2±34.6	-0.2±34.8
3	1	2	2.1±41.3	-2.1±42.9	0.4±16.5	-0.4±16.6
4	1	2	3.5±51.7	-3.5±54.8	1.0±12.4	-1.0±12.6
5	1	2	–	–	3.2±15.2	-3.2±16.2
1	2	2	0.5±14.9	-0.5±15.1	0.1±29.2	-0.1±29.2
2	2	2	1.0±46.9	-1.0±47.8	0.2±33.6	-0.2±33.7
3	2	2	2.1±53.5	-2.1±55.6	0.5±18.2	-0.5±18.4
4	2	2	3.9±74.4	-3.9±79.4	1.6±15.2	-1.6±15.7
5	2	2	–	–	5.1±17.8	-5.0±19.7

Table 77: The percentage difference between the electron and muon fake factors with and without a scaling of the diboson prompt MC of 15% as defined in Eq. 9.3. The fake factors are determined for the signal region where the isolation for μ_2 is `FixedCutLoose` and the isolation for e is `Gradient`. The fake factor bins applied are listed in Tab. 22 and 25 for the muon and electron fake factors respectively.